

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:06:08 ; Search time 40 seconds  
(without alignments)  
940.272 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDFGALPEINSARMYAGP.....SGVLVPPRPYVMPHSPAAG 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1949	100.0	391	2 B70608	probable PPE prote
2	1656.5	85.0	396	2 H70741	probable PPE prote
3	1583	81.2	393	2 C70568	probable PPE prote
4	779.5	40.0	393	2 G70929	probable PPE prote
5	751	38.5	409	2 A70932	probable PPE prote
6	737.5	37.8	403	2 H70931	probable PPE prote
7	736	37.8	423	2 B70931	probable PPE prote
8	731	37.5	421	2 H87056	PPE-family protein
9	705	36.2	408	2 G70925	probable PPE prote
10	702	36.0	391	2 B70625	probable PPE prote
11	702	36.0	413	2 F70560	probable PPE prote
12	688.5	35.3	463	2 C70931	probable PPE prote
13	682.5	35.0	468	2 G70932	probable PPE prote
14	675	34.6	380	2 A70846	probable PPE prote
15	668.5	34.3	394	2 G70881	probable PPE prote
16	667.5	34.2	385	2 H70503	probable PPE prote
17	635	32.6	350	2 H70929	probable PPE prote
18	633.5	32.5	365	2 H70929	probable PPE prote
19	615.5	31.6	402	2 A70882	probable PPE prote
20	603	30.9	423	2 C70882	probable PPE prote
21	597.5	30.7	391	2 D70922	probable PPE prote
22	589	30.2	406	2 G70675	probable PPE prote
23	588	30.2	394	2 A70504	probable PPE prote
24	575.5	29.5	391	2 A70663	probable PPE prote
25	495.5	25.4	3300	2 D70575	probable PPE prote
26	473	24.3	180	2 G70834	probable PPE prote
27	452.5	23.2	3716	2 G70969	probable PPE prote
28	448.5	23.0	580	2 G70570	probable PPE prote
29	444	22.8	346	2 H70874	probable PPE prote

## ALIGNMENTS

### RESULT 1

B70608

Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: B70608

R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70608

A>Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-391 <CO>

A:Cross-references: GB:Z93777; GB:AL123456; NID:g3261726; PIDN:CAB07839.1; PID:e311073; I

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PPE

Query Match	100.0%	Score	1949	DB	2	Length	391
Best Local Similarity	100.0%	Pred. No.	3.5e-108				
Matches	391	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MVDFGALPEINSARMYAGPGSASILVAAAQMWDSVASDLFSAASAFQSVWGLTVGSMIG	60				
DB	1	MVDFGALPEINSARMYAGPGSASILVAAAQMWDSVASDLFSAASAFQSVWGLTVGSMIG	60				
QY	61	SSAGLMVAAAAPYVAVMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI	120				
DB	61	SSAGLMVAAAAPYVAVMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI	120				
QY	121	LIATNLLGNTPTAIIVNEAEYGEEMWQAQAAAFVAAATATATATLLPPEAPEMPTSAAG	180				
DB	121	LIATNLLGNTPTAIIVNEAEYGEEMWQAQAAAFVAAATATATATLLPPEAPEMPTSAAG	180				
QY	181	LLEQAAAVVEESDSTAANQLMNNVPAQLQAQPTQGTTPSSKLGGLWKTYSRHSPTSN	240				
DB	181	LLEQAAAVVEESDSTAANQLMNNVPAQLQAQPTQGTTPSSKLGGLWKTYSRHSPTSN	240				
QY	241	MVSMANNHMTNSGVSMNTLSSNMLKGFAPAAAQVQTAQNGVRAMSLGSSGSSG	300				
DB	241	MVSMANNHMTNSGVSMNTLSSNMLKGFAPAAAQVQTAQNGVRAMSLGSSGSSG	300				
QY	301	LGCGVAAALGRAASVGSLSVFPQAAANQAQVTPAARALPLTSLTSAAREGQCMGLGLPV	360				
DB	301	LGCGVAAALGRAASVGSLSVFPQAAANQAQVTPAARALPLTSLTSAAREGQCMGLGLPV	360				
QY	361	QMGARACGGSLGVLRVPPRPYVMPHSPAAG	391				
DB	361	QMGARACGGSLGVLRVPPRPYVMPHSPAAG	391				



Db 353 AAEAPGALFGEALSSLAGRALAGTAVRSGAARV 388

RESULT 6

H70931

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: H70931

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Sqaers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: H70931

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residue: 1-403 <COL>

A:Cross-references: GB:AL022021; GB:AL123456; NID:G3250699; PIDN:CAA17728.1; PID:e125461

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PPE

Query Match 37.8%; Score 737.5; DB 2; Length 403;

Best Local Similarity 42.6%; Pred. No. 1,4e-36;

Matches 181; Conservative 60; Mismatches 123; Indels 61; Gaps 12;

Qy 2 VDFGALPPEINARMYAGPGSALVAAQWDSVASDLFSAASAFQSVVWGLTVGWSIGS 61

Db 5 LDFATLPPPEINARMYSGAGSAPMLAAASAWHGLSAELRASALSYSSVLSTLTGEEWHGP 64

Qy 62 SAGLMVAAASPYVAVWMSVTAGCAELTAQVRVAAAYETAYGLTVPPPPVIAENRAELMIL 121

Db 65 ASASMTAAAPYVAVWMSVTAVRAGAGQAABAAAAYEAAFAATVPPPVIEANRAQLMAL 124

Qy 122 IATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLIPFEAPEMTSAGGL 181

Db 125 IATNVLQNPAPIAATEAQYAEWMSQDAMVGYAGASAAAT-QLTPFPETVQTTNWSGL 183

Qy 182 LEQAAAVEASDTAA-----NOLMNVVPOALQOLAQPTQCT-----TPSS 222

Db 184 AACSAAIAHATGASAGAGAQTTLSQLIAAIPSVLQGLSSSTAATFASGPGSLGIVGSGSS 243

Qy 223 KLGSLWKTIVSPHRSPISNMVSMANNHMSWTNSGVN-TNTLS---SMLKGFAPAAAAQAV 278

Db 244 WLDKLMALLDPN-----SNFWNTIASGGLFSPNTIAPFLGGLGVAADAAGDV 293

Qy 279 QTAQNGVRAMSSIGSSL-----GSSGLGGVVAANLGRAASVGSLSVPOAANAQAVTP 333

Db 294 LGEATSG-----GLGGALVAPLGSAGGLGGTVAAGLGNAAATVGTLSVPPSWTAAAPLSP 348

Qy 334 AARAL---PLTSLTSAABERGPGMLGGIPVGOMGARAGGSLGVLRVPE---RPVVMPH 366

Db 349 LGSALGTEPMVAPPVPAVAG---MPGMFFGTMGQGGF-----RAVFQYGRPRPVAR 398

Qy 387 SPRAAG 391

Db 399 PPAAG 403

RESULT 7

B70931

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: B70931

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: B70925  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-423 <COL>  
 A;Cross-references: GB:AL020201; GB:AL123456; NID:g3250699; PIDN:CAA17722.1; PID:e125461  
 A;Experimental source: strain H37Rv  
 C;Genetics:  
 A;Gene: PPE

Query Match 37.8%; Score 736; DB 2; Length 423;  
 Best Local Similarity 41.8%; Pred. No. 1.8e-36;  
 Matches 182; Conservative 50; Mismatches 123; Indels 80; Gaps 12;

QY 2 VDFGALPEINARMYAGPGSASLVAAQWDSVASFSAASAFQSVVWGLTVGWSIGS 61  
 DB 1 MDFGLLPPEINSGRMVYGPFGPGLAATAWDGLVHATAAGYASELSALT-GAKSGP 59

QY 62 SAGLMVAAASPYYAWMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 121  
 DB 60 SSTSMASAAAPYYAWMSATAVHAELAGAQARLAIAAYEAAFAATVPPVIAANRAQLMVL 119

QY 122 IATNLLGONTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPFEAPMTSAGG 181  
 DB 120 IATNIFGONTPAIMTEAQWEMWMAQDAAMFYGAASSATA-SRMATFTEPPOTTNHHGL 178

QY 182 LEQAAVEEASDTAAAN-----QLMNNYFQALQLOAQT-----OGTTP-----S 221  
 DB 179 GAQSSVAQTAATAAGGNLQSAFPLLAVPRALQGLALPTASQASATPOMVTDLGNLS 238

QY 222 SKLGLWKTSPHRSPISNMVMANNHMTNSGVSMNTLSSMLKGFAPAAAQVQTA 281  
 DB 239 TFLGG--AVTGPYTFP-----GVLPSPGVPYLLGIQSVL-----V 271

QY 282 AQNGVAMSLGS-----SLGSSGLGGG-VAANLGRAASVGSLS 319  
 DB 272 TONGGVSAALLGKIGKPKITGALAPLAEFALHTPLGSEGLGGSVSAGIGRAGLVGKLS 331

QY 320 VPQWMAANAQVTPAARALPLTSLS---AERGPQGMGLGVPGONGARAGGGLSVLR 376  
 DB 332 VPQGWTVAAPEIESPAAALQATRLAAPIAATDAGALLGMAISLGLAGRAAGSTG--- 388

QY 377 VPPRPVMPHSPAAG 391  
 DB 389 ---HPIGSAAAPVG 400

RESULT 8  
 H87056  
 PPE-family protein [imported] - Mycobacterium leprae  
 C;Species: Mycobacterium leprae  
 C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C;Accession: H87056  
 R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho-  
 eam, M.A.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
 Nature 409, 1007-1011, 2001  
 A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sul-  
 A;Title: Massive gene decay in the leprosy bacillus.  
 A;Reference number: A86909; MUID:21128732; PMID:11234002  
 A;Accession: H87056  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-421 <STO>  
 A;Cross-references: GB:AL450380; NID:g13093150; PIDN:CAC31563.1; GSPDB:GNO0147  
 C;Genetics:  
 A;Gene: ML182

Query Match 37.5%; Score 731; DB 2; Length 421;  
 Best Local Similarity 41.4%; Pred. No. 3.5e-36;  
 Matches 174; Conservative 55; Mismatches 160; Indels 34; Gaps 8;

QY 1 MVDFGALPEINARMYAGPGSASLVAAQWDSVASFSAASAFQSVVWGLTVGWSIG 60  
 DB 1 MFDFAALSPETNRMVYLPGLGSSPILTAATAAWVYLAKELTAAAGLQSAVEAL-LTTFFEG 59

QY 61 SSAGLMVAAASPYYAWMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
 DB 60 ESAAALAEVTPPYBKWLTONAASAEELTATQUTVAANAYETAFMTVPLPVFVFNRAQCL 119

QY 121 LIATNLLGONTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPFEAPMTSAGG 180  
 DB 120 LIMSNIFGONSTAIKAEAEYETWMIQDAAMTSYQASVLEAVGATKATFAPPLGVNVEG 179

QY 181 L-----LEQAAVEEASDTAAANQLMNN-----VQALQO-----LAQP 214  
 DB 180 LAQEVVEEWEVVEEVEEVEEAEQAIQSAALQDAVNEGMEATVVPVQVQVNVVDVATP 239

QY 215 TQGTTPSSKLG--LWKTSPHRSPISNMVMANNHMTNSGVSMNTLSSMLKGFAPA 272  
 DB 240 QTAVPDDSSAAAPQLWGGFAQHLSPINDTISMINNHAGMANAGLSLVNGSGAMKSLAP- 298

QY 273 AAAQAVQTAQNGVAMSLGSLSGSSGLGGVAAANLGRAASVGSLSVPQWMAANAQV 332  
 DB 299 TTTRAASAFKAMGSAVQSTGRGLLGSSSGHVTAQLGRAASISGLRVPQTWTTTASQV 358

QY 333 PAARALPLTSLSAERGPQGM--GGLPVGOM--GARAGGSLSVLRVPRPVMPHSP 388  
 DB 359 AATRALSPARVAVATESAPLLGGLPMAPMPVCGSGTGTVNTALRLOPRAFVMPRNP 418

QY 389 AAG 391  
 DB 419 AAG 421

RESULT 9  
 G70925  
 Probable PPE protein - Mycobacterium tuberculosis (strain H37Rv)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C;Accession: G70925  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: G70925  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-408 <COL>  
 A;Cross-references: GB:274024; GB:AL123456; NID:g3250700; PIDN:CAA98377.1; PID:e1301025;  
 A;Experimental source: strain H37Rv  
 C;Genetics:  
 A;Gene: PPE

Query Match 36.2%; Score 705; DB 2; Length 408;  
 Best Local Similarity 41.8%; Pred. No. 1.1e-34;  
 Matches 182; Conservative 47; Mismatches 134; Indels 72; Gaps 14;

QY 2 VDFGALPEINARMYAGPGSASLVAAQWDSVASFSAASAFQSVVWGLTVGWSIGS 61  
 DB 1 MDFGLVPPEINSGRMVYGPFGPGLAATAWDGLVHATAAGYASELSALT-GAKSGP 60

QY 62 SAGLMVAAASPYYAWMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 121  
 DB 61 AAASWVAATVYVAVLSATAGQAEQAGQARAAAAYELAFAMTVPPVIVVNRALLVAL 120

QY 122 IATNLLGONTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPMTSAGG 181  
 DB 121 VATNFFGONTPAIAATEAQYAEWMAQDAAMFYAGSAAIAT-ELTFTTAPVTTSPAL 179

QY 182 LEQAAA-----VEEASDTAAANQLMNN-----VP--QALQO-LAQPTQGTTPSSKL--- 224

```
Db 180 AGQAAATVSTVPLATTAAVPLQLOLSSTSLIPWYSLAQWLAEHLGLTDPNRMIV 239
QY 225 -----CGLWKTSPHRSPTSNMVMANNMTNSGVSMTNTLSSMLKGFAPAAA 275
Db 240 RLLGISYFDEGL-----LQFEASLAQAIPGTPGAG--DSGSSVLDSWGPTIFA 287
QY 276 QAVCTAAQNGVRWSSL--CGSSLGS-----SGLGGVAAVLGRAASVGSLS 319
Db 288 -----GPRASPVSAGGAVGGVQTPOPYWYWDRESIGSVAALGKSSAGSLS 338
QY 320 VPOWAAANQAVTPAARALP--LTSLSAERGPQMLGGLPVQGMGARAGGSLGVLR 376
Db 339 VPDWAAARAWAPAWPLPDGDDVTALRGTAENA---LKGFFMASAGSTGGGF--VHK 393
QY 377 VPPRPYMPHSPAAG 391
Db 394 YGFLAWMQRPPEAG 408

RESULT 10
B70625
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70625
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70625
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-391 <COL>
A:Cross-references: GB:Z92539; GB:AL123456; NID:G3261714; PIDN:CAB06873.1; PID:e304546;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: PPE
```

```
Query Match 36.0%; Score 702; DB 2; Length 391;
Best Local Similarity 42.3%; Pred. No. 1.6e-34;
Matches 166; Conservative 62; Mismatches 150; Indels 14; Gaps 7;

QY 2 VDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIGS 61
Db 1 MDFGALPPEINSARMYAGAGPMWAGAAWGLAAELGTTAASVSVTRLTTSWMPG 60
QY 62 SAGLMVAASPYVAMSVTAGQAEHTAAQVRVAAAYETAYGLTVPVPPVIAENRAELMIL 121
Db 61 ASMAVAAAPYLAWLTYTEAAHAGSQAMASAAAYEAYAMTVPEVVAANRALLAAL 120
QY 122 IATNLLGONTPTAIVNEAEYGENWAOQAAAFYAAATATATATLTPFEAPEMTSAGGL 181
Db 121 VATNVLGINTPTAIVMATEALYAEWMAQDALAMYAAASG-AAQMLQLPSPPQTTNPGGL 179
QY 182 LEQAAAVEEASDTAAANQ-----LMNVPOALCQLAQPTGGTTPPSKLGGLWKTSPHRS 236
Db 180 AAGSAVGSNAATAVNVQSVADLISLPNAVSGLASPTSVLSDTGLSGIADIDALLA 239
QY 237 P--ISNVSVANNHMS--MTNSGVSMTNTLSSMLKGFAPAAAQAVTAAQNGVRWMSLSG 293
Db 240 TPEVANIINSVNTAAWYVNAIPTAIFLANALNSGAPVAIEGAIEAAEG---AASAAA 296
QY 294 SFLGSSGLGGVAAVNGRAASVGSLSVPOWAAANQAVTPAARALPLTSLTSAERGPQ 353
Db 297 AGLADSVTPAGLGASGEATLVRSLSPAAWSTAAATAGATALEGGSGWTVAABEA-QP 355
QY 354 MLGGLPVQGMGARAGGSLGVLRVPPRPYMP 385
Db 356 VTGMWP-GWASAAKGTGAYAGPRYGFKPTMP 386.
```

## RESULT 11

F70560

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C&gt;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: F70560

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70560

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-413 &lt;COL&gt;

A:Cross-references: GB:Z95436; GB:AL123456; NID:G3261770; PIDN:CAB08826.1; PID:e316565; I

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PPE

## Query Match

Best Local Similarity 39.1%; Pred. No. 1.7e-34;

Matches 168; Conservative 61; Mismatches 145; Indels 56; Gaps 7;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60

Db 1 MLDFPALPPEVNSALMYAGPGSGPMLAAAAWEALAAELQTTASTYDALITGLADGPWQG 60

QY 61 SSAGLMVAASPYVAMSVTAGQAEHTAAQVRVAAAYETAYGLTVPVPPVIAENRAELMI 120

Db 61 SSAASVMAAATPQVAMLRSTAGQAEQAGSQAVAAASAYEAFFATVPPEIAANRALLMA 120

QY 121 LIATNLLGONTPTAIVNEAEYGENWAOQAAAFYAAATATATATLTPFEAPEMTSAGG 180

Db 121 LLATNFLGONTAAATAETAQYAEWMAQDAAMVYAGASAAAT-QLSPFNPAQITINPAG 179

QY 181 LLEQAAVEEASDTAAANQLMNVNPOALCQLAQPTGGTTPPSKLGGLWKTSPHRSISN 240

Db 180 LASQAASVQAVSGVGAANAQAALTDIPKAL-----FGLSGIFNPPMLTDLKG 226

QY 241 MVSMAHNMSTNSGVSMTNTLSSMLKGFAPAA---AAQAVQT----- 280

Db 227 ALGLTCHTWSSDGGSLIVGGVIGDFVQGVTSABELDASVMDTFKXVSPARLMVTQPKD 286

QY 281 -----AAQNGVRAMSSLSGSSLSGSGVGAANLGRAASVGSLSVPOA 323

Db 287 YFGLAHDLPKWASEGAKAAGEAKALPAAPVPAIPSAGL-SGVAGAVGQAASVGGGLKVPV 345

QY 324 WAAANQAVTPAARALPLTSLTSAERGPQMLGGLPVQGMGARAGGSLGVL--RVPPRP 381

Db 346 WTATTPAASPVALAASNLGAAGAAABEGSTHAFGGMLP--MGSGAGRAFNFAPRYGFKP 403

QY 382 YMPHSPAAG 391

Db 404 TVIAQPPAGG 413

## RESULT 12

C70931

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C&gt;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: C70931

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome



Db 290 GAASVGNVTLASVGRANSIGQLSVPPSWAAPSTRPVSAISPAGLTLTLPCTDVAEHGMPG- 348  
QY 354 MLGGLPVQGMGARAGGLSGVL-RVPPRPVYVMPHSPAAG 391  
Db 349 -VPGVPV-----AAGRASGVLPYGYRLTVMAHPPAAG 380

RESULT 15  
G70881  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
R;Accession: G70881  
C;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A;Authors: Squires, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: G70881  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-394 <COL>  
A;Cross-references: GB:AL008967; GB:AL123456; NID:g3261491; PIDN:CAA15564.1; PID:e117389  
A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: PPE

Query Match 34.3%; Score 668.5; DB 2; Length 394;  
Best Local Similarity 40.8%; Pred. No. 1.6e-32;  
Matches 161; Conservative 61; Mismatches 156; Indels 17; Gaps 7;

QY 2 VDFGALPPEINSRMVGSGSLVAAAQWDSVASDLFSAASAFOSVVMGLTVGSGWGS 61  
Db 1 MDFGALPPEINSTRMVGAGAPLMAAGATNGLAVELSTASSVESVIMQLTTEQWLGP 60  
QY 62 SAGLMVAASPYVAMSVTGAQBELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMIL 121  
Db 61 ASMSVVAAPYLAWLTYTHESAHAAQAQMAAAAFEAFAFMVTPPAEVANRALLAAL 120  
QY 122 IATNLLGQNTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLTFPEAPEMTSAGGL 181  
Db 121 VATNVLGQNTPTAIVMATEAHYGEWMAQDALAMYGAASSA-AAGRLNPLITPSQTANNAGL 179  
QY 182 LEQAAAVEASDTAAHQ-----LMNVFQALQQLAQPTGTTTPSSKIGLWKTVP--H 234  
Db 180 AGQAAAVSHAAASTVQQVGLGLSLNLPNAVMGFASPLTSAADAAGLGGITQDIEELG 239  
QY 235 RSPISNVSMANHMSTNGSVMTNTLSSMLKGFAPAAAQAVQTAQNGV---RAMGS 291  
Db 240 ITFQNAINGAVN-----TTAFVMAIPNAVFLGHAFALNPATVTAADAVPAAAAAAG 295  
QY 292 LGSLSGSSGLGG-GVAANLGRASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERG 350  
Db 296 LAHTVTPVGVGASLTASLGEASVGSLSVPAGWSTAAPAMTSGTTALEGSGWAVPEEAG 355  
QY 351 PGQMLGGLPVQGMGARAGGLSGVLVPRPPYVMP 385  
Db 356 PVAAMPGN-AGISGAAGAGAYAGPYGPKPIVMP 389

Search completed: August 25, 2004, 03:13:46  
Job time : 41 secs

This Page Blank (copy)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:05:03 ; Search time 118 Seconds

(without alignments)  
1045.488 Million cell updates/sec

Title: us-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDFGALPEINSARMYAGP.....SGVLVPPRPVYMPHPSPAAG 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mmc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriopl:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1949	100.0	391	16	C05298 mycobacteri
2	1920.5	98.5	390	16	Q7U0E9
3	1779.5	91.3	396	16	Q7U071
4	1648	84.6	393	16	Q8VIZ3
5	1583	81.2	393	16	C06341
6	1583	81.2	393	16	Q7UWE5
7	779.5	40.0	393	16	Q7TZJ3
8	779.5	40.0	411	16	O53939
9	754	38.7	410	2	Q99Q11
10	751	38.5	409	16	O53957
11	749	38.4	409	16	Q7TZH7
12	740.5	38.0	399	16	Q7TZH8
13	737.5	37.8	403	16	O53956
14	736	37.8	423	16	O53950
15	736	37.8	423	16	Q7TZI4
16	731	37.5	421	16	Q9Z5K0

17	702	36.0	391	16	P96362
18	702	36.0	391	16	Q7U0E5
19	702	36.0	413	16	O06386
20	683.5	35.1	694	16	Q8VJW0
21	682.5	35.0	468	16	O53958
22	675	34.6	380	16	P95190
23	675	34.6	380	16	Q7TX66
24	670.5	34.4	385	16	Q7TZ87
25	669.5	34.4	394	16	Q7TX5
26	668.5	34.3	462	16	O33310
27	667.5	34.2	385	16	O33204
28	666.5	34.2	385	16	Q8VJZ0
29	641	32.9	364	16	Q7TZ35
30	637	32.7	350	16	Q7TZ32
31	635	32.6	363	16	O53940
32	633.5	32.5	365	16	O86373
33	629.5	32.3	405	16	Q8VJW5
34	626.5	32.1	381	16	Q7TX67
35	618.5	31.7	397	2	Q9AGF0
36	615.5	31.6	382	16	Q7TXX3
37	615.5	31.6	402	16	O33312
38	610	31.3	443	16	Q8VKL9
39	604	31.0	443	16	Q7U242
40	603	30.9	423	16	Q7U114
41	603	30.9	426	16	O05907
42	597.5	30.7	391	16	O05798
43	594.5	30.5	391	16	Q7TX76
44	589	30.2	406	16	P71869
45	589	30.2	406	16	Q7TW99

#### ALIGNMENTS

RESULT 1: 100%  
O05298 PRELIMINARY; PRT; 391 AA.  
ID C05298  
AC O05298;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein (PPE family protein).  
GN RV1196 OR MTC1364.08 OR MT1234.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37Rv;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.B. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares K.,  
RA Sulston J.B., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RL complete genome sequence.";  
RN [2]  
RP Nature 393:537-544(1998).  
SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RL laboratory strains.";  
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

P96362	mycobacteri
Q7U0E5	mycobacteri
O06386	mycobacteri
Q8VJW0	mycobacteri
O53958	mycobacteri
P95190	mycobacteri
Q7TX66	mycobacteri
Q7TZ87	mycobacteri
Q7TX5	mycobacteri
O33310	mycobacteri
O33204	mycobacteri
Q8VJZ0	mycobacteri
Q7TZ35	mycobacteri
Q7TZ32	mycobacteri
O53940	mycobacteri
O86373	mycobacteri
Q8VJW5	mycobacteri
Q7TX67	mycobacteri
Q9AGF0	mycobacteri
Q7XX3	mycobacteri
O33312	mycobacteri
Q8VKL9	mycobacteri
Q7U242	mycobacteri
Q7U114	mycobacteri
O05907	mycobacteri
O05798	mycobacteri
Q7TX76	mycobacteri
P71869	mycobacteri
Q7TW99	mycobacteri



```

QY 181 LLEQAAVEEASDTAAANLNNVPOALQOLAQTOGTTSSKLGGLWKTVPSPHRSPI 240
DB 181 LLEQAAVEEASDTAAANLNNVPOALQOLAQTOGTTSSKLGGLWKTVPSPHRSPI 240
QY 241 MVSMAHHMNTSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRMSS-----IGSSL 296
DB 241 IVSMLNHHVMTSGVSMNTLSSMLKGFAP-AAAQAQVTAQAQNGVRMSSIGSSL 299
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRGQMLG 356
DB 300 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRGQMLG 359
QY 357 GLPVGQMGARAG--GGLSGVLRVPPRPVPMHPSPAAAG 391
DB 360 GLPLGLTNSGGGFGVSNALRMPRPAYVMPVPAAG 396

RESULT 4
Q8VIZ3
ID Q8VIZ3 PRELIMINARY; PRT; 393 AA.
AC Q8VIZ3;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE PBE family protein.
GN MT3582.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
RA Danchin A., Uitterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007161; AAK47941.1; -.
DR TIGR; MT3582; -.
DR GO; GO:0016020; C.membrane; IEA.
DR GO; GO:0008233; F.peptidase activity; IEA.
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.
DR InterPro; IPR000030; Microbac_PPE.
DR InterPro; IPR000508; Peptidase_S26.
DR Pfam; PF00823; PPE; 1.
DR PROSITE; PS00501; SPASE_I.1; 1.
SQ SEQUENCE 393 AA; 39658 MW; 86F0B67798855511 CRC64;

Query Match 84.6%; Score 1648; DB 16; Length 393;
Best Local Similarity 84.8%; Pred. No. 2e-85;
Matches 335; Conservative 20; Mismatches 34; Indels 6; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVGLTVGSWIG 60
DB 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVGLTVGSWIG 60
QY 61 SSAGLMVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
DB 61 SSAGLMVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
DB 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
QY 181 LLEQAAVEEASDTAAANLNNVPOALQOLAQTOGTTSSKLGGLWKTVPSPHRSPI 240
DB 181 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
QY 241 LLEQAAVEEASDTAAANLNNVPOALQOLAQTOGTTSSKLGGLWKTVPSPHRSPI 296
DB 241 LLEQAAVEEASDTAAANLNNVPOALQOLAQTOGTTSSKLGGLWKTVPSPHRSPI 296
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRGQMLG 356

```

```

DB 241 IVSMLNHHVMTSGVSMNTLSSMLKGFAP-AAAQAQVTAQAQNGVRMSSIGSSL 299
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRGQMLG 356
DB 300 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRGQMLG 359
QY 357 GLPVGQMGARAGGGLSGVLRVPPRPVPMHPSPAAAG 391
DB 360 GLPLGH-SVNAAGSINNLRVPAAYAIPTPAAG 393

RESULT 5
Q06341
ID Q06341 PRELIMINARY; PRT; 393 AA.
AC Q06341;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein RV3478.
GN RV3478 OR MTCY13E12.31.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Hailroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
DR EMBL; Z95390; CAB08702.1; -.
DR PIR; C70568; C70568.
DR TubercuList; RV3478; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
SQ SEQUENCE 393 AA; 39413 MW; AF4C20C95DAE7DD4 CRC64;

Query Match 81.2%; Score 1583; DB 16; Length 393;
Best Local Similarity 81.5%; Pred. No. 9.3e-82;
Matches 322; Conservative 20; Mismatches 47; Indels 6; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVGLTVGSWIG 60
DB 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVGLTVGSWIG 60
QY 61 SSAGLMVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
DB 61 SSAGLMVAASPYVAMSVTAGAELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
DB 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 180
QY 181 LLEQAAVEEASDTAAANLNNVPOALQOLAQTOGTTSSKLGGLWKTVPSPHRSPI 240
DB 181 LLEQAAVEEASDTAAANLNNVPOALQOLAQTOGTTSSKLGGLWKTVPSPHRSPI 240
QY 241 MVSMAHHMNTSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRMSS-----IGSSL 296
DB 241 VSSIANNHSMGTVSMNTLSSMLKGLAP-AAAQAQVTAQAQNGVRMSSIGSSL 299
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRGQMLG 356

```

```

Db 300 GSSGLGAGVAANLGRAASVGSLSVPPFAWAANQVTPAARALPLTSLTSAQAQAPGHMLG 359
QY 357 GLPVQGVGAGAGGSLGSLVLRPPRPVYMPHSPAAG 391
Db 360 GLPLGH-SVNAGSGINNLRVPAARAIPAIPRTPAAG 393

RESULT 6
Q7TWFS PRELIMINARY; PRT; 393 AA.
AC Q7TWFS;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE PE family protein.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97; PubMed=12788972;
RX MEDLINE=22709107;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RL EMBL; BX248346; CAD95692.1; -.
KW Complete proteome.
SQ SEQUENCE 393 AA; 38572 MW; 41944B6B547A8A80 CRC64;

Query Match 40.0%; Score 779.5; DB 16; Length 393;
Best Local Similarity 43.7%; Pred. No. 26-36;
Matches 179; Conservative 65; Mismatches 129; Indels 37; Gaps 10;

QY 2 VDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVMGLTVGSMIGS 61
Db 1 MDFGALPPEVNSVEMYAGPGSAPVAAASAWNGLAELSSAATGYETVITQLSSEGWLGP 60
QY 62 SAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 121
Db 61 ASAAVAAVAPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120
QY 122 IATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPMTSAGGL 181
Db 121 ISTNVFGNTSAIAAAEAQYGEWMAQDAAMFGYAAATATATATLLPFEAPMTSAGGL 179
QY 182 LEQAAVVEASDTAAA--NQLMNVPOALQALQAOPTGTPSSKLGGLWK----- 229
Db 180 GTQAAVATAAGTAQSTLTITEMITGLPNALQSLTSPQLQSS--NGPLSLWQLFLGTNPFT 238
QY 230 -----TVSPHRSPISNMVSMANNHVMNTNSGVSMNTLSSMLKGFAPAAAQAQVATA 282
Db 239 SISALLTDQYASFFNYTEGLPYFSIGMGNFIOAAKTL--GLIGSAAPAAVA----- 292
QY 283 QNGVRANSSLGSSSLGSSGLGGVAANLGRAASVGSLSVPOAWAA--ANQAVTPAARALPLT 341
Db 293 GDAAKGLPGLGMLG---GGPVAAGLGNAAASVGLSVPPVWMSGPLPGSVTPGAAPLPVS 348
QY 342 SLTSAAGRGCPQMLGGLPVGMQGARAGGSLGSLVLRPPRPVYMPHSPAAG 391
Db 349 TVSAAPAAAPGSLGGLPL-----AGAGGAGAGP--RYGFRPTVMARPPFAG 393

RESULT 8
OS3939 PRELIMINARY; PRT; 411 AA.
AC OS3939;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-2003 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PBE-family protein.
GN RV1789 OR MT1838 OR MTV049.11.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv.
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

```

```

Db 300 GSSGLGAGVAANLGRAASVGSLSVPPFAWAANQVTPAARALPLTSLTSAQAQAPGHMLG 359
QY 357 GLPVQGVGAGAGGSLGSLVLRPPRPVYMPHSPAAG 391
Db 360 GLPLGH-SVNAGSGINNLRVPAARAIPAIPRTPAAG 393

RESULT 6
Q7TWFS PRELIMINARY; PRT; 393 AA.
AC Q7TWFS;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE PE family protein.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97; PubMed=12788972;
RX MEDLINE=22709107;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RL EMBL; BX248346; CAD95692.1; -.
KW Complete proteome.
SQ SEQUENCE 393 AA; 38572 MW; 41944B6B547A8A80 CRC64;

Query Match 81.2%; Score 1583; DB 16; Length 393;
Best Local Similarity 81.5%; Pred. No. 9.3e-82;
Matches 322; Conservative 20; Mismatches 47; Indels 6; Gaps 3;

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVMGLTVGSMIG 60
Db 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVMGLTVGSMIG 60
QY 61 SAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120
Db 61 SAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPMTSAGGL 180
Db 121 LTATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPMTSAGGL 180
QY 181 LLEQAAVVEASDTAAANQLMNVPOALQALQAOPTGTPSSKLGGLWKTVSPHRSPISN 240
Db 181 LLEQAAVVEASDTAAANQLMNVPOALQALQAOPTGTPSSKLGGLWKTVSPHRSPISN 240
QY 241 MVMANNHSMNTNSGVSMNTLSSMLKGFAPAAAQAQVATAQNGVRANSS---LGSSL 296
Db 241 VSSIANHSMNTNSGVSMNTLSSMLKGFAPAAAQAQVATAQNGVRANSS---LGSSL 299
QY 297 GSSGLGAGVAANLGRAASVGSLSVPPFAWAANQVTPAARALPLTSLTSAQAQAPGHMLG 356
Db 300 GSSGLGAGVAANLGRAASVGSLSVPPFAWAANQVTPAARALPLTSLTSAQAQAPGHMLG 359
QY 357 GLPVQGVGAGAGGSLGSLVLRPPRPVYMPHSPAAG 391
Db 360 GLPLGH-SVNAGSGINNLRVPAARAIPAIPRTPAAG 393

RESULT 7
Q7TWJ3 PRELIMINARY; PRT; 393 AA.
AC Q7TWJ3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

```

```

RA Gordon S.V., Biglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022021; CAA17711.1; ALT_INIT.
DR EMBL; AB007043; AAK46108.1; -.
DR PIR; G70929; G70929.
DR TIGR; M1838; -.
DR TubercuList; Rv1789; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 411 AA; 40557 MW; 03C90B5E0590B7DA CRC64;

Query Match 40.0%; Score 779.5; DB 16; Length 411;
Best Local Similarity 43.7%; Pred. No. 2.1e-36;
Matches 179; Conservative 65; Mismatches 129; Indels 37; Gaps 10;

QY 2 VDFGALPPEINSARMYAGPGSALVAAAQWDSVSLFSAASAFQSVVMGLTVGSWIGS 61
DB 19 MDFGALPPEVNSVRMYAGPGSAPWVAASAWNGLAELSSAAATGYETVITQLSSGWLGP 78
QY 62 SAGLMVAASPYVAMSVTAGQAEHTAAQVRVAAAYETAYGLTVPPVIAENRAELMI 121
DB 79 ASAAAEAVAPYVAMSVAAQAQAEATQARAAAFAFAATVPPVIAENRASLMOL 138
QY 122 IATNLGQNTPAIVNEAEYGEWMAQDAAMFYAAATATATATLPPFEAPEMTSAGL 181
DB 139 ISTNVFGQNTSAIAAEAYGEWMAQDSAAAYAGSSASASA-VTFPSTPQIANPTAQ 197
QY 182 LEQAAVEEASDTAAA--NOLMNNVPAQLOLAQPTQGTTPSSKLGGLWK----- 229
DB 198 GTQAAAVATAAGTAQSTLTEMITGLFNALQSLTSPLLQSS-NGPLSLWQLLFTGFNFFT 256
QY 230 -----TVSPHSPISNNVSMANNHMTNSGVSMTNTLSMLKGFAPAAAQAQVATAA 282
DB 257 SISALLTDQPIASFFYNTGELPYFSGMGNFIQSAKIL-GLIGSAAAPAAVA----- 310
QY 283 QNGVRAMSSISGLSSGLGGGVAANLGRAASVGSLSVFOAWAA-ANQAVTFAARALPLT 341
DB 311 GDAAKGLPLGGMLG---GGPVAAGLGNAAVSGKLSVPPVWMSGPLPGSVTPGAAPLPVS 366
QY 342 SLTSAERGPQMLGGLPVGQMGARAGGGLSGVLRYPPPRPYMPHSPAAG 391
DB 367 TVSAAPAEAPGSLGLGLPL----AGAGGAGAGP-RYGFRPTVNRPPFAG 411

RESULT 9
Q99Q11 PRELIMINARY; PRT; 410 AA.
AC Q99Q11;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
GN Rv1808-like protein.

```

```

OS Mycobacterium microti.
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1806;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=myc 94-2272, and OV254;
RA Sivadon V., Heym B., Mazancourt P., Gaillard J.-L.L.;
RT "PPE Rv1808 orthologue of Mycobacterium microti."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335180; AAK20894.1; -.
DR EMBL; AF335179; AAK20893.1; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
SQ SEQUENCE 410 AA; 40044 MW; 3B9DA3174655A5EA CRC64;

Query Match 38.7%; Score 754; DB 2; Length 410;
Best Local Similarity 44.5%; Pred. No. 5.7e-35;
Matches 177; Conservative 51; Mismatches 140; Indels 30; Gaps 10;

QY 1 MVDGALPPEINSARMYAGPGSALVAAAQWDSVSLFSAASAFQSVVMGLTVGSWIG 60
DB 1 MLDFGALPPEINSGRMYAGPGSGPLAAAWDAALAAELYSAAASYGSTIEGLTVAPMWG 60
QY 61 SSAGLMVAASPYVAMSVTAGQAEHTAAQVRVAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 PSSITMAAAVAPYVAMSVTAGQAEHTAAQVRVAAAYETAYGLTVPPVIAENRAELMI 120
QY 121 LIATNLGQNTPAIVNEAEYGEWMAQDAAMFYAAATATATATLPPFEAPEMTSAGG 180
DB 121 LVATNIFGQNTPAIVNEAEYGEWMAQDAAMFYAAATATATATLPPFEAPEMTSAGG 179
QY 181 LLEQAAVEEASDTAAA-----NOLMNNVPAQLOLAQPTQGTTPSSKLGGLWKTVSP 233
DB 180 TAQSAVVAQAAGAAASDITQAQLSILSLPSTLQSLA--TTATATSAAG--WDTV-- 233
QY 234 HRSPTISNNVSMANNHMTNSGVS---SMTNTLSMLKGFAPAAAQAQVATAAQNQVRAMS 290
DB 234 -LQSITITLANLTGPYSITGLGAIPTGGWMLTFGQIL-GLAQNAPGVAAALLGPXAAAGALS 291
QY 291 SIGSSLGS-----SGLGGVAAANLGRAASVGSLSVFOAWAAANQAVTFAARALPLTSLTS 345
DB 292 PLAPLRGGYIADITPLGGGATGGIARATYVGSLSVPOGWAEEAPVPMRAVASVLPGTGAAP 351
QY 346 A-AERGPQMLGGLPVGQMGARAGGGL-----SGVLRV 377
DB 352 ALAAEAPCALFGEMALSSLAGFALAGTAVRSAGAARV 389

RESULT 10
O53957 PRELIMINARY; PRT; 409 AA.
AC O53957;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PPE-family protein.
GN Rv1808 OR MT1856.1 OR MTV049.30.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

```

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RL complete genome sequence";  
 RN Nature 393:537-544 (1998).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RL laboratory strains";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL020201; CAI1729.1; -;  
 DR EMBL; AE007044; AAK46129.1; ALT\_INIT.  
 DR FIC; A70932; A70932.  
 DR TIGR; MT1856.1; -;  
 DR TubercuList; Rv1808; -;  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 409 AA; 39917 MW; 1E15202BACF36379 CRC64;

Query Match 38.5%; Score 751; DB 16; Length 409;  
 Best Local Similarity 44.7%; Pred. No. 8.4e-35;  
 Matches 177; Conservative 53; Mismatches 138; Indels 28; Gaps 10;  
 QY 2 VDFGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGWSIGS 61  
 DB 1 MDFGALPPEINSGRMVAGPGSGPLLAALAAWDAALAEALYSAAASYGTIEGLTVAPWMPG 60  
 QY 62 SAGLMVAASAPYVAMSVTAGOELTAQVRVAAAYETAYGLTVPPVIAENRAELMIL 121  
 DB 61 SSITMAAAYVAVVAVISVTAGQEAQAKIAGVYETAFATVPPVIEANRALLMSL 120  
 QY 122 IATNLGQNTPAIAVNEAEYGENWQAQDAAMFGYAAATATATATLFFPEAPEMTSAGGL 181  
 DB 121 VATNIFQNTPAIAATEAHYAEMWQAQDAAMYGAGSSATA-SQLAFPSPPQTNPST 179  
 QY 182 LEQAAVVEASDTAAA-----NOLMNVPOALQLOAQTOGTTSPSKLGLWKTUSPH 234  
 DB 180 AQSVAVVAQAAGAAASDITAQLSQLISLPLSTLQSLA--TTATATSASAG--WDTV--- 232  
 QY 235 RSPISNMVSMANNHMTNSGV-----MTNTLSMLKGFAPAAAQAQVOTAAQNGVRAM 289  
 DB 233 LOSITITLANLTGYSIIIGLAIPEGWMLTFGQILGLAQNAPGVALLGPKAAGALSPL 292  
 QY 290 SSL-GSSLGS-SGLGGVAAANLGRAASVGSLSVPOQAWAANQAVTPAARALPLTSITSA- 346  
 DB 293 APLRGVIGDITPLGGGATGIARAIYVGSLSVPOGWAEEAAPVMRAVASVLEFGTGAAPAL 352  
 QY 347 AERGPQMLGGLPVGOWGARAGGL-----SGVLRV 377  
 DB 353 AAAPGALFEMALSSLAGRALAGTAVRSGAGARV 388

RESULT 11  
 QY7TZH7  
 ID QY7TZH7 PRELIMINARY; PRT; 409 AA.  
 AC QY7TZH7  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DE PPE family protein.  
 GN PPE32 OR MB1837.  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 CX NCBI\_TaxID=1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,  
 RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RT "The complete genome sequence of Mycobacterium bovis";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
 DR EMBL; BX248340; CAD94540.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 409 AA; 39931 MW; D57892628B131A9E CRC64;  
 Query Match 38.4%; Score 749; DB 16; Length 409;  
 Best Local Similarity 44.3%; Pred. No. 1.1e-34;  
 Matches 176; Conservative 51; Mismatches 140; Indels 30; Gaps 10;  
 QY 2 VDFGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGWSIGS 61  
 DB 1 MDFGALPPEINSGRMVAGPGSGPLLAALAAWDAALAEALYSAAASYGTIEGLTVAPWMPG 60  
 QY 62 SAGLMVAASAPYVAMSVTAGOELTAQVRVAAAYETAYGLTVPPVIAENRAELMIL 121  
 DB 61 SSITMAAAYVAVVAVISVTAGQEAQAKIAGVYETAFATVPPVIEANRALLMSL 120  
 QY 122 IATNLGQNTPAIAVNEAEYGENWQAQDAAMFGYAAATATATATLFFPEAPEMTSAGGL 181  
 DB 121 VATNIFQNTPAIAATEAHYAEMWQAQDAAMYGAGSSATA-SQLAFPSPPQTNPST 179  
 QY 182 LEQAAVVEASDTAAA-----NOLMNVPOALQLOAQTOGTTSPSKLGLWKTUSPH 234  
 DB 180 AQSVAVVAQAAGAAASDITAQLSQLISLPLSTLQSLA--TTATATSASAG--WDTV--- 232  
 QY 235 RSPISNMVSMANNHMTNSGV-----MTNTLSMLKGFAPAAAQAQVOTAAQNGVRAMSS 291  
 DB 233 LOSITITLANLTGYSIIIGLAIPEGWMLTFGQILGLAQNAPGVALLGPKAAGALSPL 291  
 QY 292 LGSLSGS-----SGLGGVAAANLGRAASVGSLSVPOQAWAANQAVTPAARALPLTSITSA 346  
 DB 292 LAPLRGGVIADITPLGGGATGIARAIYVGSLSVPOGWAEEAAPVMRAVASVLEFGTGAAPA 351  
 QY 347 AERGPQMLGGLPVGOWGARAGGL-----SGVLRV 377  
 DB 352 LAAPGALFEMALSSLAGRALAGTAVRSGAGARV 388

RESULT 12  
 QY7TZH8  
 ID QY7TZH8 PRELIMINARY; PRT; 399 AA.  
 AC QY7TZH8  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DE PPE family protein.  
 GN PPE31 OR MB1836.  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 CX NCBI\_TaxID=1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,  
 RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RT "The complete genome sequence of Mycobacterium bovis";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
 DR EMBL; BX248340; CAD94539.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 399 AA; 38840 MW; 1A0F4377318E74F2 CRC64;  
 Query Match 38.0%; Score 740.5; DB 16; Length 399;



RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Unwayan L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL020201; CA017722.1; -;  
 DR EMBL; AE007044; AAK46122.1; -;  
 DR PIR; B70931; B70931.  
 DR TIGR; MT1850; -;  
 DR Tuberculist; RV1801; -;  
 DR InterPro; IPR00030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 423 AA; 41477 MW; 26E52CC271FBBF57 CRC64;  
 Query Match 37.8%; Score 736; DB 16; Length 423;  
 Best Local Similarity 41.8%; Pred. No. 6.2e-34;  
 Matches 182; Conservative 50; Mismatches 123; Indels 80; Gaps 12;  
 QY 2 VDFGALPPEINARMYAGPGSASLVAAAQWDSVADLFSAAAFQSVVWGLTVGWSIGS 61  
 DB 1 MDFGLLPPEINSGRMVTPGPGPMLAAATAWDGLAVELHATAAGYASELSALT-GAWSGP 59  
 QY 62 SAGLMVAASPYVAVMSVTAGQAEHTAAQVRAAAVETAYGLTVPPVIAENRAELMIL 121  
 DB 60 SSTSMASAAAPYVAVMSATVHAELAGAQARLAIAAYEAFAATVPVPIAANRAQLMVL 119  
 QY 122 IATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGGL 181  
 DB 120 IATNIFGQNTPAIMWTEAQYEMWMAQDAAMFYAGSSATA-SRMTAFTEPPQTTHGQL 178  
 QY 182 LEQAAVEEASDTAAAN-----QLMNVVPAQLQLOAQT-----QGTFP-----S 221  
 DB 179 GAQSSAVAQTAATAAGGNLQSAFPQLLSAVPRALQGLALPTASQSASATPQWVTDLGNLS 238  
 QY 222 SKLGLMKTVPSPHRSPTISNMVSMANNHMTNSGVSMTLSSMLKGFAPAAAAQAVQTA 281  
 DB 239 TFLGG--AVTGPYTFP-----GVLPPSGVPYLLGIQSVL-----V 271  
 QY 282 AQNGVRAMSSILGS-----SLGSSGLGGG--VAANLGRAASVGSLS 319  
 DB 272 TONGQGVSAALLKIGKGPITGALAPLAEFALHTPILGSEGLGGSVSAGIGRAGLVGKLS 331  
 QY 320 VPQAWAANQAVTPAARALPLTSLTS---AAERPGQMGLGGLPVGOMGARAGGSLGVLR 376  
 DB 332 VPQGWTVAAEIPSPRAALQATRLAAPIAATDAGALLGGMALSLAGRAAAGSTG--- 388  
 QY 377 VPPRPYVMPHSPAAAG 391  
 DB 389 ---HPIGSAAAPAVG 400

RESULT 15

Q7TZ14 PRELIMINARY; PRT; 423 AA.  
 ID Q7TZ14  
 AC Q7TZ14  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE PPE family protein.  
 GN PPE29 OR MB1829.

OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Biglieri K., Camus J.-C., Medina N., Mansoor H.,  
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
 RA Harris B., Akin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RT "The complete genome sequence of Mycobacterium bovis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
 DR EMBL; BX248340; CAD94532.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 423 AA; 41477 MW; 26E52CC271FBBF57 CRC64;  
 Query Match 37.8%; Score 736; DB 16; Length 423;  
 Best Local Similarity 41.8%; Pred. No. 6.2e-34;  
 Matches 182; Conservative 50; Mismatches 123; Indels 80; Gaps 12;  
 QY 2 VDFGALPPEINARMYAGPGSASLVAAAQWDSVADLFSAAAFQSVVWGLTVGWSIGS 61  
 DB 1 MDFGLLPPEINSGRMVTPGPGPMLAAATAWDGLAVELHATAAGYASELSALT-GAWSGP 59  
 QY 62 SAGLMVAASPYVAVMSVTAGQAEHTAAQVRAAAVETAYGLTVPPVIAENRAELMIL 121  
 DB 60 SSTSMASAAAPYVAVMSATVHAELAGAQARLAIAAYEAFAATVPVPIAANRAQLMVL 119  
 QY 122 IATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGGL 181  
 DB 120 IATNIFGQNTPAIMWTEAQYEMWMAQDAAMFYAGSSATA-SRMTAFTEPPQTTHGQL 178  
 QY 182 LEQAAVEEASDTAAAN-----QLMNVVPAQLQLOAQT-----QGTFP-----S 221  
 DB 179 GAQSSAVAQTAATAAGGNLQSAFPQLLSAVPRALQGLALPTASQSASATPQWVTDLGNLS 238  
 QY 222 SKLGLMKTVPSPHRSPTISNMVSMANNHMTNSGVSMTLSSMLKGFAPAAAAQAVQTA 281  
 DB 239 TFLGG--AVTGPYTFP-----GVLPPSGVPYLLGIQSVL-----V 271  
 QY 282 AQNGVRAMSSILGS-----SLGSSGLGGG--VAANLGRAASVGSLS 319  
 DB 272 TONGQGVSAALLKIGKGPITGALAPLAEFALHTPILGSEGLGGSVSAGIGRAGLVGKLS 331  
 QY 320 VPQAWAANQAVTPAARALPLTSLTS---AAERPGQMGLGGLPVGOMGARAGGSLGVLR 376  
 DB 332 VPQGWTVAAEIPSPRAALQATRLAAPIAATDAGALLGGMALSLAGRAAAGSTG--- 388  
 QY 377 VPPRPYVMPHSPAAAG 391  
 DB 389 ---HPIGSAAAPAVG 400

Search completed: August 25, 2004, 03:13:02  
 Job time : 121 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:57:58 ; Search time 25 Seconds  
(without alignments)  
814.377 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949  
Sequence: 1 MVDFFGALPPEINSARMYAGP.....SGVLRVPPRPYVMPHSPAAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1556.5	85.0	366	1 YD61_MYCTU	Q1031 mycobacteri
2	705	36.2	408	1 YS92_MYCTU	Q10813 mycobacteri
3	688.5	35.3	463	1 Y102_MYCTU	O53951 mycobacteri
4	444	22.8	487	1 Y442_MYCTU	P42611 mycobacteri
5	426.5	21.9	443	1 Y878_MYCTU	Q10540 mycobacteri
6	418	21.4	678	1 YF48_MYCTU	Q10778 mycobacteri
7	364.5	18.7	408	1 SRA_MYCLE	Q07297 mycobacteri
8	324.5	16.6	463	1 Y096_MYCTU	Q10892 mycobacteri
9	324	16.6	434	1 YU18_MYCTU	P31500 mycobacteri
10	321.5	16.5	435	1 YU21_MYCTU	O53268 mycobacteri
11	233.5	12.0	178	1 YX29_MYCTU	O06246 mycobacteri
12	217.5	11.2	176	1 YX25_MYCTU	O50703 mycobacteri
13	153.5	7.9	860	1 ELS_MOUSE	P54320 mus musculu
14	151.5	7.8	232	1 YX23_MYCTU	Q50702 mycobacteri
15	150	7.7	864	1 ELS_RAT	Q99372 rattus norv
16	143.5	7.4	730	1 ELS_HUMAN	P15502 homo sapien
17	139.5	7.2	881	1 PRY3_YEAST	P47033 saccharomyc
18	135	6.9	907	1 A180_HUMAN	O60641 homo sapien
19	133.5	6.8	825	1 ICP0_HSV2H	P28284 herpes simp
20	133.5	6.8	2432	1 Y43R_IRV6	P18305 chilo iride
21	132	6.8	1120	1 STFR_ECOLI	P78072 escherichia
22	132	6.8	2090	1 N214_HUMAN	P35558 homo sapien
23	131	6.7	836	1 YG26_BPM15	O05233 mycobacteri
24	131	6.7	1150	1 APMU_PIG	P12021 sus scrofa
25	129	6.6	1783	1 RAA3_CHLRE	O9fec4 chlamydomon
26	128	6.5	779	1 SRP_DROME	P52172 drosophila
27	127.5	6.5	790	1 SRP_NOTCO	P24856 notothenia
28	126	6.5	354	1 YAU6_SCHPO	Q10169 schizosacch
29	125.5	6.4	1211	1 BUN2_DROME	Q24523 drosophila
30	125	6.4	577	1 CST2_HUMAN	P33240 homo sapien
31	124	6.4	2090	1 HFC1_MESAU	P51611 mesocricetu
32	123.5	6.3	677	1 YI36_MYCTU	O50597 mycobacteri
33	123	6.3	432	1 YF10_MYCTU	P71789 mycobacteri

34	123	6.3	1140	1 YN96_YEAST	Q04893 saccharomyc
35	122	6.3	394	1 HYF1_ALCEU	P45805 aicalligenes
36	121.5	6.2	1025	1 SLAP_CAUCR	P35828 caulobacter
37	121.5	6.2	1845	1 Z236_HUMAN	Q9u136 homo sapien
38	121	6.2	1199	1 P121_RAT	P52591 rattus norv
39	119.5	6.1	635	1 HMLA_DROME	P10105 drosophila
40	119	6.1	915	1 A180_RAT	O05140 rattus norv
41	118.5	6.1	580	1 EXPR_XANCP	P23314 xanthomonas
42	118.5	6.1	2038	1 FSH_DROME	P31709 drosophila
43	118	6.1	1508	1 BCSX_XANAC	P89938 xanthomonas
44	117	6.0	444	1 Y808_CHLPN	Q92798 chlamydia p
45	117	6.0	652	1 PICA_HUMAN	Q13492 homo sapien

ALIGNMENTS

```
RESULT 1
YD61_MYCTU
ID YD61_MYCTU STANDARD; PRT; 396 AA.
AC Q11031;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Hypothetical PPE-family protein RV1361C/MT1406.
GN RV1361C OR MT1406 OR MTCY02B10.25C.
OS Mycobacterium tuberculosis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98295987; PubMed=9634230;
RX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z75555; CA99966.1; -.
CC EMBL: AEC07013; AAK45669.1; -.
CC PIR: H70741; H70741.
CC TIGR: MT1406; -.
CC -----
CC InterPro: IPR000030; Microbac_PPE.
```

```

DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 158 159 TA -> AT (IN REF. 2).
SQ SEQUENCE 396 AA; 40015 MW; 6AFAE0D7B5F668D0 CRC64;

Query Match 85.0%; Score 1656.5; DB 1; Length 396;
Best Local Similarity 85.1%; Pred. No. 1.3e-89;
Matches 338; Conservative 18; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSWIG 60
DB 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSWIG 60
QY 61 SSAGLMVAASPYVAMSVTAGQBELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 SSAGLMVAASPYVAMSVTAGQBELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
QY 121 LIATNLGQNTPALVAVNEAEGEMWAOAAAGYAAATATATATATLLPPEEAPMTSAGG 180
DB 121 LIATNLGQNTPALVAVNEAEGEMWAOAAAGYAAATATATATATLLPPEEAPMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNVPQALQLOAQTOGTTPSSKLGGLWKTVPSPHRSPIGN 240
DB 181 LLEQAAVEEASDTAAANQLMNVPQALQLOAQTOGTTPSSKLGGLWKTVPSPHRSPIGN 240
QY 241 MVSMAHNMSTNSGVSMNTLSMLKGFAPAPAAAQVQTAQNGVRAMSS-----LGSSL 296
DB 241 IVSMLNHNVTNSGVSMNTLSMLKGFAP-AAAQAVETAAQNGVQVQVMSLSGLSGL 299
QY 297 GSSGLGGVGAANLGRAASVGSLSVPQAWAANAQVTPAARALPLTSLTAAEAGPGQMLG 356
DB 300 GSSGLGAGVAANLGRAASVGSLSVPQAWAANAQVTPAARALPLTSLTAAEAGPGHMLG 359
QY 357 GLPVGQMGARAG--GGLSGVLVPPRPPVMPHSPAAG 391
DB 360 GLFLGQLTSGGGGGSVNALRPPRAYVMPRVEAAG 396

RESULT 2
YS92_MYCTU STANDARD; PRT; 408 AA.
AC Q10813;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein RV2892c/MT2959/Wb2916c.
GN RV2892C OR MT2959 OR MTC174.23C OR MB2916C.
OS Mycobacterium tuberculosis, and
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger J., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=2206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,

```

```

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutchy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -!- SIMILARITY: Belongs to the mycobacterial pPE family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z74024; CAA98377.1; -.
CC EMBL; AE007119; AAK47285.1; -.
CC EMBL; BX248344; CAD96603.1; -.
CC PIR; G70925; G70925.
CC TIGR; MT2959; -.
CC TubercuList; RV2892c; -.
CC InterPro; IPR000030; Microbac_PPE.
CC Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 56 76 POTENTIAL.
SQ SEQUENCE 408 AA; 41469 MW; 3E3D1F20D7827199 CRC64;

```

```

Query Match 36.2%; Score 705; DB 1; Length 408;
Best Local Similarity 41.8%; Pred. No. 2.7e-34;
Matches 182; Conservative 47; Mismatches 134; Indels 72; Gaps 14;

QY 2 VDFGALPPEINSARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSWIGS 61
DB 1 MDFGVLPEINSGRMYAGPGSGPMWAAAADWSLAELGLAAGGYRLAISLTGAYWAGP 60
QY 62 SAGLMVAASPYVAMSVTAGQBELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 121
DB 61 AAASWVAATPYVAVLWLSATAGQAEQAGQARAAAAYELAFAMTVPVVVAVNRALVAL 120
QY 122 IATNLGQNTPALVAVNEAEGEMWAOAAAGYAAATATATATLLPPEEAPMTSAGGL 181
DB 121 VATNFFQNTPALVAVNEAEGEMWAOAAAGYAAATAT-ELTFTTAAPTTSPAAL 179
QY 182 LEQAAA-----VEEASDTAAANQLMN-VP--QALQQ-LAQPTQGTTPSKL--- 224
DB 180 AGQAAATVSVTPVPLATTAAVPLLQQLSSLSLIPWYSALQWLAEENLLGLTPNRTIV 239
QY 225 -----GGLWKTVPSPHRSPISNMYSMAHNMSTNSGVSMNTLSMLKGFAPAAA 275
DB 240 RLLGISVDEGL-----IQFEASLQAQAIPTPGGAG--DSGSSVDSWGPTIFA 287
QY 276 QAVQTAAQNGVRAMSSL--GSSLSG-----SGLGGVAAANLGRAASVGSLS 319
DB 288 -----GPRASPVAGGAVGVQTPQPYWYTWALDRESIGGSVSAALCKGSSAGSL 338
QY 320 VPQAWAANAQVTPAARALP---LTSLSAERPGQVGLGGLPVGQMGARAGGLSVLR 376
DB 339 VPPDWAARAWANPAARLPGDDVTLARGTAENA---LLRGFPMAAGQSTGGF--VHK 393
QY 377 VPPRPPVMPHSPAAG 391
DB 394 YGRLAVMQRPFFAG 408

```

RESULT 3  
Y102 MYCTU STANDARD; PRT; 463 AA.  
AC O53951;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical PPE-family protein RV1851/MD1830.  
GN RV1802 OR MT1851 OR MT049.24 OR MB1830.  
OS Mycobacterium tuberculosis, and  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773, 1765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekai F.,  
Badock K., Basham D., Brown D., Chillingworth T., Connor R.,  
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,  
Rutter S., Seeger K., Skelton S., Squares S., Taylor J.E.,  
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence."  
RL Nature 393:537-544 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=22206494; PubMed=12218036;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
laboratory strains."  
RL J. Bacteriol. 184:5479-5490 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.bovis; STRAIN=AF2122/97;  
RX MEDLINE=22709107; PubMed=12788972;  
RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,  
Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
RT "The complete genome sequence of Mycobacterium bovis."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
CC -I- SIMILARITY: Belongs to the mycobacterial PPE family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; AL022021; CAAL7723.1;  
DR ENBL; AEO07044; AAK46123.1;  
DR ENBL; BX248340; CAD94533.1;  
DR PIR; C70931; C70931.  
DR TIGR; MT1851;  
DR InterCellist; RV1802;  
DR InterPro; IPR00030; Microbac\_PPE.  
DR Pfam; PF00823; PPE; 1.  
KW Hypothetical protein; Complete proteome.  
FT CONFLICT 401 401 S -> L (IN REF. 2).

SQ SEQUENCE 463 AA; 46021 MW; EE64828BF09FA551 CRC64;  
Query Match 35.3%; Score 688.5; DB 1; Length 463;  
Best Local Similarity 42.8%; Pred. No. 2.8e-33;  
Matches 174; Conservative 48; Mismatches 136; Indels 49; Gaps 11;  
QY 2 VDFGALPPEINARMYAGPGSGASLVAAAQMWDSVASDFSAASAFQSWGLTVGWSIGS 61  
DB 1 MDFGVLPPEINSGRMVYAGPGSGPMLAAAAAWDGLATELQSTADYGVISVLT-GVWSGQ 59  
QY 62 SAGLMVAAAASPYVAMVVTAGQBELTAAQVRYAAAAAYETAYGLTVPPPTVAENPAELMIL 121  
DB 60 SSGTMAAAAAPYVAMVGTATLALAREAAAQAASAAAAYEAAFAATVPPPVVAANRAELAVL 119  
QY 122 IATNLGONTPAIAVNEAEYGEVMAQDAAMFGVAAAATATATATATLLPFEPAPETSAGGL 181  
DB 120 AATNIFGONTGAIAAEAARYAEVMAQDAAMVYAGSSVAT-QVTFAPAPPTTNAAGL 178  
QY 182 LEQAAVVEASDTAAANQLMNVPAQLQQAQPTGTTTSSKLGGLMKTVS--PHRSP1- 238  
DB 179 ATQGVAVAAQVAGSAGN-ARSLVSEVLEFLA--TAGTNYNKTVASLMNAVTVGVYASSVY 235  
QY 239 -----SNMYSMANNHSMNTNSGVSMNTLSSMLKGFAPAAAAQAQVTAQAQNGVRA 288  
DB 236 NSMLGLGFAESKVMVLPANDTVISTIFGVQVQKFFNPTFPNPLIFK----- 283  
QY 289 MSSLGSSSLG-----SSGLGG---GVAANLGEAASVGSLSVPQAAANQAVTPAARALPL 340  
DB 284 -SALGAGLGLSSALSSGLGSTAPAISAGASQAGSVGMSVPPSWAAATPAIRTVAAVFSS 342  
QY 341 TSLTS--AAERGPQOML-----GGLPVQOMGARAGGGLSGVLRV 377  
DB 343 TGLQAVPAAAISEGLLSQMALASVAGGALGAAARATGGFLGGRV 389  
RESULT 4  
Y442 MYCTU STANDARD; PRT; 487 AA.  
ID Y442 MYCTU STANDARD; PRT; 487 AA.  
AC P42611; O53727;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical PPE-family protein RV0442C/MT0458.  
GN RV0442C OR MT0458 OR MT037.06C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Erdmann;  
RX MEDLINE=87137260; PubMed=3029018;  
RA Shinnick T.M.;  
RT "The 65-kilodalton antigen of Mycobacterium tuberculosis."  
RL J. Bacteriol. 169:1080-1088 (1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37Rv;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekai F.,  
Badock K., Basham D., Brown D., Chillingworth T., Connor R.,  
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,  
Rutter S., Seeger K., Skelton S., Squares S., Taylor J.E.,  
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence."  
RL Nature 393:537-544 (1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;

```

RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; M15467; AAA88235.1; ALT INIT.
DR EMBL; AL021932; CAAL17399.1; -.
DR EMBL; AE006948; AAK44681.1; -.
DR EMBL; C70830; C70830.
DR TIGR; MT0458; -.
DR TubercuList; RV0442c; -.
DR InterPro; IPR00030; Microbac PPE.
DR InterPro; IPR002989; Mycobac_Pentapep.
DR Pfam; PF01469; Pentapeptide_2; 5.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 40 40 E -> K (IN REF. 2).
FT CONFLICT 96 96 I -> T (IN REF. 1).
FT CONFLICT 211 211 G -> GNNIG (IN REF. 1).
SQ SEQUENCE 487 AA; 47247 MW; 97234D5B316C8C7F CRC64;

Query Match 22.88; Score 444; DB 1; Length 487;
Best Local Similarity 32.18; Pred. No. 4.8e-19; Indels 58; Gaps 12;
Matches 135; Conservative 48; Mismatches 180;

Qy 4 FGALPPEINSARMYAGPGSASLVAAQMDVSDLFSAAPQSVVWGLTVGSGWIGSSA 63
Db 6 FAWLPPEINSALMFAGPGSGPLIAAATAWCELAEELLASIASIGSVTSELTSGLMIGPSA 65
Qy 64 GLMVAASPPVAMKSVTAGQAEITAAQVRVAAAYETATGLVPPPIAENRAELMILIA 123
Db 66 AAMVAVATQYLAWLSTAAQAQEAQAQAATAFAEALAAATVQPAVVAANRGLMQLAA 125
Qy 124 TNLGQNTPTAIAVNEAEYGEWMAQDAAMPGYAAATATATATLLPPEAPEMTSAG--- 179
Db 126 TNNFGQNPALMDVEAAAYEQMWALDVANMAGYHFDASAAVAQLAPQOV--LRNLGIDIG 183
Qy 180 --GLLEQAAVEASDTAAANQNMNNVPAALQOLAQPTGTTPTSSKLG-----GLW 228
Db 184 KNGQINLFGNTGSGNIGNNNIGNNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTG 243
Qy 229 KTVS-----PHRSPISNMYSMANNHSMNNSGVNTNTLSMLKGFAPAAAQAQV 279
Db 244 NTGSGNIGFQITGDHQMGFCGNSGSGN-IGFNGSNTGNGVLFNS----- 287
Qy 280 TAAQNGVRAMSSLGSGSLGSGGVAANLGRASVGSLSVPAWAAANQVTPAARALP 339
Db 288 GSGNIGIGSGSLNSIGTSGT---INAGLG---SAGSLNT-SFWNAGNQNAALGSAAGS 340
Qy 340 LTSLTSAERGPQM-----LGLPVGQMGARAG--GGLSGVLR--VPPRPVMPHSPAA 390
Db 341 EAAVLSAGVATGNTAALSSGILLASALGSTGLGHLANVLSGLTNPVAPASAPV 400
Qy 391 G 391
Db 401 G 401

```

```

Y878_MYCTU STANDARD; PRT; 443 AA.
ID Y878_MYCTU Q10540;
AC Q10540;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein RV0878c/MT0901.
GN RV0878C OR MT0901 OR MTCY31.06C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; Z73101; CAA97385.1; -.
DR EMBL; AE006977; AAK45143.1; ALT_INIT.
DR TIGR; C70780; C70780.
DR TubercuList; RV0878c; -.
DR InterPro; IPR00030; Microbac PPE.
DR InterPro; IPR002989; Mycobac_Pentapep.
DR Pfam; PF01469; Pentapeptide_2; 4.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Repeat; Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT DOMAIN 64 73 POLY-ALA.
FT DOMAIN 81 115 ALA-RICH.
FT DOMAIN 231 270 4 X 10 AA APPROXIMATE REPEATS.
SQ SEQUENCE 443 AA; 43592 MW; C58BEC607F0675E2 CRC64;

Query Match 21.98; Score 426.5; DB 1; Length 443;
Best Local Similarity 31.28; Pred. No. 4.5e-18;
Matches 125; Conservative 56; Mismatches 132; Indels 87; Gaps 14;

```

```

QY 2 VDFGALPPEINSMYAGPSSASLVAAQWDSVADLPSAASAFQSVVWGLTVG---SW 58
DB 1 MNFWLPPPEVNSARIYAGAPAPMLAAVAVMDGLAAELGVAASLSLISGLTAGPGSAW 60
QY 59 IGSSAGLVAAASPYVAMSVTGAQELTAAQVRVAAAYETAYGLTVPVPIAENRAEL 118
DB 61 QGPAAAMAAAAPYLSWLNATARAAGAAAGAKAAAVYEAARAATAHPALVAANRNL 120
QY 119 MILIATNLGONTPAIADVNEAYGEMWQDAADAMGYAAATATATATLPPPEAEPMTS 178
DB 121 LSLVLNLFQONLPAIATAEASYEQWQDVAAVGVYHGGAATVASQLTPWQQ----- 173
QY 179 GGLLEQAAAEASDTPAAANQMLMNNVPOALQLOAQPTQGTTPSSKLGGLWKTVPSPHRSPI 238
DB 174 --LLSVLPFPVTAAPAGV-----GVFAA--LAIPALGV--ENIG-----V 208
QY 239 SNMV-----SWANNHMTNSG-----VSMTH-----TLSSMLKGFAPAAAAQAV 278
DB 209 GNPLGIGNNNGVSGNTGDNFYGIGNGNANLGNIGNANLGNIGNANLGNIGNANLGN 268
QY 279 QTAAGQVRAAMSLSGS-SLGSSSLGGGVAA--NLGRAASVGSLSVFPQAAANQAVTFAA 335
DB 269 NTNFGSGNAGFLNIGSGNEGSLNFGNAGDDNTG-----WNSGD----- 309
QY 336 RALPLTSLTSAABRGQOMLGGJ--PVGQVAGARAGGLSG 373
DB 310 -----TNTGGFNSGDLNTGIGSPVTOGVANSFGFNTG 341

RESULT 6
YF48_MYCTU STANDARD; PRT; 678 AA.
ID YF48_MYCTU
AC 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1989 (Rel. 38, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein RV1548C/MT1599.
GN RV1548C OR MT1599 OR MTCV48.17.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RN Nature 393:537-544(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RC MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RN Bacteriol 184:5479-5490(2002).
RL [2]
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; Z74020; CA98335.1;
DR EMBL; AE007026; AAK45866.1; ALT_INIT.
DR PIR; A70762; A70762.
DR TIGR; MT1599;
DR TubercuList; RV1548C;
DR InterPro; IPR000303; Microbac_PPE.
DR InterPro; IPR002989; Mycobac_Pentapep.
DR Pfam; PF01469; Pentapeptide_2; 11.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 180 200 POTENTIAL.
FT CONFLICT 258 258 D -> G (IN REF. 2).
SQ SEQUENCE 678 AA; 66736 MW; 209F1593D52533A2 CRC64;

Query Match 21.4%; Score 418; DB 1; Length 678;
Best Local Similarity 31.9%; Pred. No. 2.2e-17;
Matches 106; Conservative 48; Mismatches 140; Indels 38; Gaps 7;

QY 2 VDFGALPPEINSMYAGPSSASLVAAQWDSVADLPSAASAFQSVVWGLTVGSMIGS 61
DB 1 MNFWLPPPEVNSARIYAGAPAPMLAAVAVMDGLAAELGVAASLSLISGLTAGPGGP 60
QY 62 SAGLMVAAAAPVVAWMSVTAGQAEITAAQVRVAAAYETAYGLTVPVPIAENRAELMIL 121
DB 61 ASAMTGVAAAYARWLTTAAQAQAQAQAQAQAQAQAQAQAQAQAQAQAQAQAQAQA 120
QY 122 IATNLLGONTPAIADVNEAYGEMWQDAADAMGYAAATATATATATATATATATATAT 180
DB 121 VASNLGQVAPAIATAEAVYEQWMAADVAAMLGHYGEASAVALSILTPFPSPSAATPGG 180
QY 181 -----LLEQAAVEEASTAAANQMLMNNVPOALQLOAQPTQGTTPSSKLGGLWKTVS 232
DB 181 AVIAGFPFLDNGNTVIGGFNLASNLGLNLGS-----FNPGSANTGSVNLGN----- 229
QY 233 PHRSPISNN--VSMANNHMTNSGVSMTNTLSSMLKGFAPAAAAQAVQTAQONVRAMS 290
DB 230 -----ANIGDLNLSGNGISVNLGGGTGDLN-----PDGSGNTGTINWGSNGISYN 276
QY 291 SLGSSLSGSLGGGVA--ANLGRAASVGSLSV 320
DB 277 LGGGNLGSYNLGSNGTGTNFG-GGNTGNLNV 307

RESULT 7
SRA_MYCLE STANDARD; PRT; 408 AA.
ID SRA_MYCLE
AC Q07297;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine-rich antigen (25L) (45 kDa protein).
DE SRA OR ML0411 OR MLC1383.14.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=9323928; PubMed=8478104;
RA Vega-Lopez F., Brooks L.A., Dockrell H.M., de Smet K.A.L.,
RA Thompson J.K., Hussain R., Stoker N.G.;
RA "Sequence and immunological characterization of a serine-rich antigen
RT from Mycobacterium leprae.";
RN Infect. Immun. 61:2145-2153(1993).
[2]
RP SEQUENCE FROM N.A.

```

QY	214	-----PTQ-----GTPPSKLGGLWKTVPSPHRSPISSNMVSMANNHMTSGVGMTNT 261
Db	240	PFETPSQSSQNDLSATSLTQQLGGL-----DSIISSASASLLTTNS--ISSST 286
QY	262	LSSMLKGFAPAAQAQVTAQNGVRAMSSILGSSGLGGGVAAN-----LGRAASVG 316
Db	287	ASSIM-----PIVASQVTELTGRSQV-AVERKMIOSSISTAVSDVAASKVAVGQVAVSG 341
QY	317	SLSVQPAWAAANQAVTPAARALP--LTSLSAAERGPQMLGGLPQMGARAGGGLSGV 374
Db	342	ALRVENWATASQPMATAHVSVPAGCSAIIITA-----VSGPLEGV 381
QY	375	LRVPRPVYMPHSPAAG 391
Db	382	IQ--PAEEVLTSVAGG 396
RESULT 8		
Y096.MVCTU		
ID	Y096.MVCTU	STANDARD; PRT; 463 AA.
AC	Q10892;	
DT	01-OCT-1996	(Rel. 34, Created)
DT	01-OCT-1996	(Rel. 34, Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)
DE	Hypothetical PPE-family protein Rv0096/MT0105.	
GN	RV0096 OR MT0105 OR MTCY251.15.	
OS	Mycobacterium tuberculosis.	
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.	
OX	NCBI_TaxID=11773;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H378V;	
RC	MEDLINE=98295987; PubMed=9634230;	
RX	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,	
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,	
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,	
RA	Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,	
RA	Hornby S., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,	
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,	
RA	Rutter S., Seeger K., Skelton S., Squares S., Squares R.,	
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;	
RT	"Deciphering the biology of Mycobacterium tuberculosis from the	
RT	complete genome sequence."	
RL	Nature 393:537-544(1998).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CDC 1551 / Oshkosh;	
RC	MEDLINE=22206494; PubMed=12218036;	
RX	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,	
RA	Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,	
RA	Kolcanov J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,	
RA	Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,	
RA	Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;	
RT	"Whole-genome comparison of Mycobacterium tuberculosis clinical and	
RT	laboratory strains."	
RL	J. Bacteriol. 184:5479-5490(2002).	
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).	
CC	-!- SIMILARITY: Belongs to the mycobacterial PPE family.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; Z74410; CAA98932.1; --	
DR	EMBL; AE006922; AAK44327.1; --	
DR	PIR; H70750; H70750.	
DR	TIGR; MT0105; --	
DR	TubercuList; RV0096; --	

QY	214	-----PTQ-----GTPPSKLGGLWKTVPSPHRSPISSNMVSMANNHMTSGVGMTNT 261
Db	240	PFETPSQSSQNDLSATSLTQQLGGL-----DSIISSASASLLTTNS--ISSST 286
QY	262	LSSMLKGFAPAAAQAQVTAQNGVRAMSSILGSSGLSGGGVAAN-----LGRAASVG 316
Db	287	ASSIM-----PIVASQVTETLGRSQV-AVERKMIOSSISTAVSDVAASKVAVGQVAVSG 341
QY	317	SLSVQPAWAAANQAVTPAARALP--LTSLSAAERGPQMLGGLPQMGARAGGGLSGV 374
Db	342	ALRVENWATASQPMATAHVSVPAGCSAIIITA-----VSGPLEGV 381
QY	375	LRVPRPVYMPHSPAAG 391
Db	382	IQ--PAEEVLTSVAGG 396
RESULT 8		
Y096.MVCTU		
ID	Y096.MVCTU	STANDARD; PRT; 463 AA.
AC	Q10892;	
DT	01-OCT-1996	(Rel. 34, Created)
DT	01-OCT-1996	(Rel. 34, Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)
DE	Hypothetical PPE-family protein Rv0096/MT0105.	
GN	RV0096 OR MT0105 OR MTCY251.15.	
OS	Mycobacterium tuberculosis.	
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC	Corynebacteriineae; Mycobacteriaceae; Mycobacterium.	
OX	NCBI_TaxID=11773;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H378V;	
RC	MEDLINE=98295987; PubMed=9634230;	
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,	
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,	
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,	
RA	Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,	
RA	Hornby S., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,	
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,	
RA	Rutter S., Seeger K., Skelton S., Squares S., Squares R.,	
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;	
RT	"Deciphering the biology of Mycobacterium tuberculosis from the	
RT	complete genome sequence."	
RL	Nature 393:537-544(1998).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CDC 1551 / Oshkosh;	
RC	MEDLINE=22206494; PubMed=12218036;	
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,	
RA	Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,	
RA	Kolcanov J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,	
RA	Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,	
RA	Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;	
RT	"Whole-genome comparison of Mycobacterium tuberculosis clinical and	
RT	laboratory strains."	
RL	J. Bacteriol. 184:5479-5490(2002).	
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).	
CC	-!- SIMILARITY: Belongs to the mycobacterial PPE family.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; Z74410; CAA98932.1; -	
DR	EMBL; AE006922; AAK44327.1; -	
DR	PIR; H70750; H70750.	
DR	TIGR; MT0105; -	
DR	TubercuList; RV0096; -	

```

DR InterPro: IPR000030; Microbac_PPE.
DR Pfam: PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 88 108 POTENTIAL.
FT TRANSMEM 122 132 POTENTIAL.
FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 276 296 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 419 439 POTENTIAL.
SQ SEQUENCE 463 AA; 46894 MW; 42D9D66A033D0DD8 CRC64;

Query Match 16.6%; Score 324.5; DB 1; Length 463;
Best Local Similarity 27.8%; Pred. No. 4e-12;
Matches 112; Conservative 54; Mismatches 144; Indels 93; Gaps 12;

QY 6 ALPPEINARMYAGPGSASLVAAQWDSVADLFSAASAFQSVYVWGLTVGSSAGL 65
Db 2 AIPPEVHSGLLGAGCGPGSLVAAQWQELSDOYALACAEGLGQVQASSWQGTAA 61
QY 66 MYAAASPYVAMSVTGAQAELETAQVRVAAAYETAYGLTVPPVIAENRAELMILIATN 125
Db 62 YVAAHGPYLAWLEQTAINSAVTAACHVAAAYCSALAAMPTPABLAANHAITHGLVLIATN 121
QY 126 LLQNTPTAVNEAEYGENWQDAAMFGYAAATATATATLLPFEAPEMTSAGLLEOA 185
Db 122 FFGINTVPTALNEADYRVRLWQLADTMAYQVADATVAVSTQTPAPPIRAPGG 176
QY 186 AAVEEASDT-----AAANQLMNVPAQLQAQPTQ-----GTPSSKL 224
Db 177 -----DAADTDLVSSIGQLIRDI---LDFIANPKYKYLEFEEQFGFSPAVTVLALNAL 229
QY 225 ---GGW---KTVSPHRSPISNMVSMANNHMTNSGVSMNTLSSMLK---GF 269
Db 230 QLYDFLWYFYASYGILLPLFPFFTP-----TSLATLALSLIHLNLPAGL 275
QY 270 AAAAAQAQVTAQNGVRAMSSSLGSLGGVAAANLGRAASVGSLSVQPAWAAANQ 329
Db 276 LPIAAA-----LPGCDQGANLAVATPATAAVP-----GGSP 308
QY 330 AVTPAARALPLTSLTSAABERGPG---QMLGGLPVG-QMGARAG 368
Db 309 PTNPAPAAPSSNSVGSASAAPGISYAVFLAPGVSSGPKAG 351

RESULT 9
YU18_MYCTU STANDARD; PRT; 434 AA.
AC P31500; O53265;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein Rv3018c/MT3098/MT3101.
GN Rv3018c OR MT3098/MT3101 OR MV012.32C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."

```

```

RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=2206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RN J. Bacteriol. 184:5479-5490 (2002).
RN [3]
RP SEQUENCE OF 160-374 FROM N.A.
RC STRAIN=Isolate 50410;
RA Pakki A.H., Dale J.W.;
RL Submitted (Apr-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC -!- CAUTION: In strain Oshkosh the gene for this protein is
CC interrupted in position 307 by an IS6110 element.
CC -!- CAUTION: Was originally (Ref.3) thought to be a dihydrofolate
CC reductase.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC in positions 294; 337 and 355.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AL021287; CA416103.1; -
DR EMBL; AE007129; AAK47427.1; ALT_SEQ.
DR EMBL; AE007129; AAK47430.1; ALT_SEQ.
DR EMBL; X59271; CA441961.1; ALT_FRAME.
DR PIR; E70857; E70857.
DR TIGR; MT3098; -.
DR TIGR; MT3101; -.
DR TubercuList; RV3018c; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1. Complete proteome.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 434 AA; 43029 MW; 41D673C4BD389DD6 CRC64;

Query Match 16.6%; Score 324; DB 1; Length 434;
Best Local Similarity 28.0%; Pred. No. 4e-12;
Matches 115; Conservative 56; Mismatches 178; Indels 62; Gaps 12;

QY 6 ALPPEINARMYAGPGSASLVAAQWDSVADLFSAASAFQSVYVWGLTVGSSAGL 65
Db 8 ASPPEVHSGALLSAGPGSLQAAGAGWSALSAAEYAAVAQELSVVVAAGVAGVQGPSAEL 67
QY 66 MYAAASPYVAMSVTGAQAELETAQVRVAAAYETAYGLTVPPVIAENRAELMILIATN 125
Db 68 FYAAVVPVYVAMLVQASADSAAAGHEAAAGYVCAEAEMPTLPBLAANHLTHAVLATN 127
QY 126 LLQNTPTAVNEAEYGENWQDAAMFGYAAATATATATLLPFEAPEMTSAGLLEOA 185
Db 128 FFGINTVPTALNEADYRVRLWQLADTMAYQVADATVAVSTQTPAPPIRAPGG 181
QY 186 AAVEEASDTAAAN-----OLMNVPAQLQAQPTQCTTPSSKLGGLWKTVSPH 234
Db 182 --ANEASNAVAATAATTPFPWHIIVFLBETFAAYQOYLSALLSELP--VANVWQLFVD 237
QY 235 ---RSPISNMVSMANNHMTNSGVSMNTLSSMLKGF-- 270
Db 238 ILGFNIIGFIITLASNAQLLTFEFAINASVAVVGLLYATAGVIDIVVWVIGNLFGVVP 297
QY 271 -----PAAAAQAVTAQNGVRAMSSSLGSLGGVAAANLGRAASV-GSLSVPOA 323

```



Db 298 GGPLGALAAAVPGVAGLAGVAGLAAL-PAVGAA--AGAPAAVGVSVAPVSGGWSPQA 354  
 QY 324 WAAANAVTPAARALPLTSLTSAERAGPGOMLGGLPVQMGARAGGGLSGV 374  
 Db 355 RLVS--AVEPAPASTSVSL--ASDRGAGAL--GF-VGTAGKESVGGQFAGL 398

RESULT 10  
 YU21 MYCTU STANDARD; PRT; 435 AA.  
 AC 053269; 053269;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical PPE-family protein RV3021C/RV3022C/MT3106.  
 GN RV3021C/RV3022C OR MT3106 OR MTV012.35C/MTV012.36C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206494; PubMed=12218036;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J.F., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolonay J., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL J. Bacteriol. 184:5479-5490 (2002).  
 CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.  
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 82.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL021287; CAA16106.1; ALT\_FRAME.  
 CC EMBL; AL021287; CAA16107.1; ALT\_FRAME.  
 CC EMBL; AS007129; AAK47435.1; -.  
 CC TIGR; MT3106; -.  
 CC Tuberculist; RV3021C; -.  
 CC Tuberculist; RV3022C; -.  
 CC InterPro; IPR000030; Microbac\_PPE.  
 CC Pfam; PF00423; PPE; 1.  
 KW Hypothetical protein, Complete proteome.  
 FT CONFLICT 299 G -> A (IN REF. 2).  
 FT CONFLICT 317 320 LAGV -> VTGL (IN REF. 2).  
 FT CONFLICT 326 L -> V (IN REF. 2).  
 SQ SEQUENCE 435 AA; 42876 MW; 3B157643EAA8484A CRC64;

Query Match 16.5%; Score 321.5; DB 1; Length 435;  
 Best Local Similarity 26.5%; Pred. No. 5.6e-12;  
 Matches 115; Conservative 53; Mismatches 159; Indels 107; Gaps 12;

QY 6 ALPPEINSARMYAGPGCSASLVAAQWDSVADSLFSAASAFQSVVWGLTVSGSWIGSSAGL 65  
 Db 8 ASPPEVHSALLSAGPGCSLQAAAAAGWSALSAAEYAAVAQELSVVVAAGVAGVWQGPSAEL 67  
 QY 66 MYAARSPYVNMVSTAGQAEITAAQVRVAAAAYETAYGLTVPPPVIAENRAELMILITATN 125  
 Db 68 FVAAYVPYVAMLVQASADSAAAAGEHEAAAGYVCALAEEMPTLPELANHLTHAVLATN 127  
 QY 126 LLGQNTPAIAVNEAREYGEWQAQDAAMFGYAA----- 157  
 Db 128 PFGINTIPALNEADYVEMVQAATVMSAYEAVGAALVATPHTGPAPVIVKPGANEASN 187  
 QY 158 ATATATATALLPPE-----APEMTSAGGLLQAAAEASDTAAANQLMNVVQALQ 209  
 Db 188 AVAAATITPPFPFGLAKFLEMAQAQAFTEVGELINKSAEAWAVGFVELITGLVNEEP--- 243  
 QY 210 QLAQPTQGTTPSSKLGLLWKTQVSPHRSFISNMVSK-----ANNH 248  
 Db 244 -----WLIV-----LTGNIDMFATVGFALGVFLVPLEFAVVLE 278  
 QY 249 MSMNNGVSMNTLSS-----MLKGFAPAAAAQAVQTAQNGVRAMSSILGSSSGL 301  
 Db 279 LAILSIGMIISNFGAIPVLGGLLALAAAVVPGVAGLAGVAGLAALPAVGAAGAP-- 336  
 QY 302 GGGVAANLGRASV-GSLSVPOMAAANQAVTPAARALPLTSLTSAERAGPGOMLGGLPV 360  
 Db 337 ----AALVGSVAPVSGGVVSPQARLVS--AVEPAPASTSVSL--ASDRGAGAL--GF-V 385

361 GQMGARAGGGLSGV 374  
 386 GTAGKESVGGQFAGL 399

RESULT 11  
 YU29 MYCTU STANDARD; PRT; 178 AA.  
 AC 06246;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical PPE-family protein RV3429/MT3533.  
 GN RV3429 OR MT3533 OR MTCY77.01.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206494; PubMed=12218036;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J.F., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolonay J., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL J. Bacteriol. 184:5479-5490 (2002).  
 CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.  
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 82.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL021287; CAA16106.1; ALT\_FRAME.  
 CC EMBL; AL021287; CAA16107.1; ALT\_FRAME.  
 CC EMBL; AS007129; AAK47435.1; -.  
 CC TIGR; MT3106; -.  
 CC Tuberculist; RV3021C; -.  
 CC Tuberculist; RV3022C; -.  
 CC InterPro; IPR000030; Microbac\_PPE.  
 CC Pfam; PF00423; PPE; 1.  
 KW Hypothetical protein, Complete proteome.  
 FT CONFLICT 299 G -> A (IN REF. 2).  
 FT CONFLICT 317 320 LAGV -> VTGL (IN REF. 2).  
 FT CONFLICT 326 L -> V (IN REF. 2).  
 SQ SEQUENCE 435 AA; 42876 MW; 3B157643EAA8484A CRC64;



RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
RL laboratory strains";  
CC J. Bacteriol. 184:5479-5490(2002).  
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z95389; CAB08678.1; -  
DR EMBL; AB007158; AAK47873.1; -  
DR PIR; C70975; C70975.  
DR TIGR; MT3533; -  
DR TubercuList; RV3429; -  
DR InterPro; IPR000030; Microbac\_PPE.  
DR Pfam; PF00823; PPE; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 178 AA; 19811 MW; 8BE1FC025ABFBEA6 CRC64;  
  
Query Match 12.0%; Score 233.5; DB 1; Length 178;  
Best Local Similarity 35.3%; Pred. No. 2.8e-07;  
Matches 65; Conservative 25; Mismatches 76; Indels 15; Gaps 3;  
  
QY 7 LPPEINARMYAGPGSASLVAAQ----MWDS---VASDLFSAASAFQSVWGLTVGWSI 59  
DB 5 IPAEYISNIIEYEGPGADSLGAAEQFLMYNSANMTAKSLTDLGELQE-----NWK 56  
QY 60 GSSAGLVAAASPVVAVMSVTAGQAEATAAQRVAAAYETAYGLTVPPVIAENRAELM 119  
DB 57 GSSSDLMADAGRYLDVTKHSQILETAVIDFLAVYETRHKKVVPATIANREEVH 116  
QY 120 ILATNLGONTFAIVNEAEYGEWMAQDAAMFGYAAATATATATATALLPFEAPETSAG 179  
DB 117 RLIASNVAGVTPAIGLDAQYQYRAQNTAVMNDYQSTARFILAVLPWQEPPIYGGG 176  
QY 180 G 180  
DB 177 G 177  
  
RESULT 12  
YI25\_MYCTU STANDARD; PRT; 176 AA.  
AC Q50703;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical PPE-family protein RV3425.  
GN RV3425 OR MTCY78.04C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
CX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Davlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J.S., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RL complete genome sequence.";  
RL Nature 393:537-544(1998).

CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z77165; CAB01031.1; -  
DR PIR; F70738; RV3425.  
DR TubercuList; RV3425; -  
DR InterPro; IPR000030; Microbac\_PPE.  
DR Pfam; PF00823; PPE; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 176 AA; 19855 MW; B8CEPZ9463B87B0 CRC64;  
  
Query Match 11.2%; Score 217.5; DB 1; Length 176;  
Best Local Similarity 33.5%; Pred. No. 2.3e-06;  
Matches 58; Conservative 31; Mismatches 83; Indels 1; Gaps 1;  
  
QY 7 LPPEINARMYAGPGSASLVAAQMWDSVASDLFSAASAFQSVWGLTVGWSIAGLM 66  
DB 5 IPAEYISNIIEYEGPGADSLFASGQRLAYSVETASLEDELDELDELDELDELDEL 63  
QY 67 VAAASPVVAVMSVTAGQAEATAAQRVAAAYETAYGLTVPPVIAENRAELMILATNL 126  
DB 64 ADAVERYLQWLSKSSQLKHAANVINGLANAYNDFRRKVVPEELAAANREERRRLIAGNV 123  
QY 127 LGONTFAIVNEAEYGEWMAQDAAMFGYAAATATATATATALLPFEAPETSAG 179  
DB 124 AGVNTPAIADLDAQYQYRAQNTAVMNDYQSTARFILAVLPWQEPPIYGGG 176  
  
RESULT 13  
ELS\_MOUSE STANDARD; PRT; 860 AA.  
AC P54320;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Elastin precursor (Tropoelastin).  
GN ELN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; Tissue=Lung;  
RX MEDLINE=95130069; PubMed=7829060;  
RA Wyder K.S., Sechler J.L., Boyd C.D., Passmore H.C.;  
RT "Use of an intron polymorphism to localize the tropoelastin gene to  
RT mouse chromosome 5 in a region of linkage conservation with human  
RT chromosome 7".  
RL Genomics 23:125-131(1994).  
CC -!- FUNCTION: Major structural protein of tissues such as aorta and  
CC nuchal ligament, which must expand rapidly and recover completely.  
CC -!- SUBUNIT: The polymeric elastin chains are cross-linked together  
CC into an extensible 3D network.  
CC -!- SUBCELLULAR LOCATION: Extracellular matrix of elastic fibers.  
CC -!- PTM: The crosslinks are made of deaminated Lys.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U08210; AAA80155.1; -

"Deciphering the biology of *Mycobacterium tuberculosis* from the



*This Page Blank (copy)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:13:09 ; Search time 127 Seconds  
(without alignments)  
967.510 Million cell updates/sec

Title: US-09-724-685-107  
Perfect score: 1949  
Sequence: 1 MYDFGALPEINARMYAGP.....SCVLEVPFRPYMHPSPAG 391

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1949	100.0	391	12	US-09-886-349A-14
2	1949	100.0	391	14	US-10-193-002-102
3	1949	100.0	391	14	US-10-084-843-107
4	1949	100.0	391	14	US-10-098-732A-14
5	1949	100.0	596	9	US-09-287-849-26
6	1949	100.0	596	12	US-09-886-349A-20
7	1949	100.0	596	14	US-10-359-460-26
8	1949	100.0	596	14	US-10-098-732A-20
9	1949	100.0	600	9	US-09-287-849-22
10	1949	100.0	600	14	US-10-359-460-22
11	1949	100.0	723	15	US-10-369-983-2
12	1949	100.0	729	12	US-09-886-349A-18
13	1949	100.0	729	14	US-10-098-732A-18
14	1949	100.0	729	15	US-10-369-983-21
15	1949	100.0	729	15	US-10-369-983-22

16	1949	100.0	813	15	US-10-369-983-15	Sequence 15, Appl
17	1949	100.0	825	15	US-10-369-983-14	Sequence 14, Appl
18	1949	100.0	875	15	US-10-369-983-13	Sequence 13, Appl
19	1949	100.0	930	14	US-10-098-732A-65	Sequence 65, Appl
20	1949	100.0	930	15	US-10-369-983-12	Sequence 12, Appl
21	1949	100.0	1010	15	US-10-369-983-4	Sequence 4, Appl
22	1949	100.0	1016	15	US-10-369-983-18	Sequence 18, Appl
23	1949	100.0	1022	15	US-10-369-983-17	Sequence 17, Appl
24	1949	100.0	1154	15	US-10-369-983-16	Sequence 16, Appl
25	1949	99.7	729	9	US-09-287-849-2	Sequence 2, Appl
26	1949	99.7	729	12	US-09-886-349A-16	Sequence 16, Appl
27	1949	99.7	729	14	US-10-359-460-2	Sequence 16, Appl
28	1949	99.7	729	14	US-10-098-732A-16	Sequence 2, Appl
29	1949	99.7	729	15	US-10-359-459-2	Sequence 8, Appl
30	1934	99.2	391	12	US-09-872-186-8	Sequence 106, App
31	1652.5	84.8	395	14	US-10-193-002-106	Sequence 111, App
32	1652.5	84.8	396	14	US-10-084-843-111	Sequence 62455, A
33	1583	81.2	393	12	US-10-282-122A-62455	Sequence 64892, A
34	1583	81.2	393	12	US-10-282-122A-64892	Sequence 104, App
35	1486.5	76.3	359	14	US-10-193-002-104	Sequence 109, App
36	1486.5	76.3	359	14	US-10-084-843-109	Sequence 8, Appl
37	1187.5	60.9	358	9	US-09-287-849-8	Sequence 8, Appl
38	1187.5	60.9	358	14	US-10-359-460-8	Sequence 12, Appl
39	1187	60.9	263	12	US-09-886-349A-12	Sequence 92, Appl
40	1187	60.9	263	14	US-10-193-002-92	Sequence 91, Appl
41	1187	60.9	263	14	US-10-084-843-91	Sequence 12, Appl
42	1187	60.9	263	14	US-10-098-732A-12	Sequence 126, App
43	765.5	39.3	400	9	US-09-073-009-126	Sequence 126, App
44	765.5	39.3	400	9	US-09-793-306-126	Sequence 62027, A
45	742.5	38.1	405	12	US-10-282-122A-62027	

ALIGNMENTS

RESULT 1

US-09-886-349A-14  
; Sequence 14, Application US/09886349A  
; Publication No. US20040086523A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Reed, Steven  
; APPLICANT: Alderson, Mark  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009070US  
; CURRENT APPLICATION NUMBER: US/09/886,349A  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 09/597,796  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: US 60/265,737  
; PRIOR FILING DATE: 2001-02-01  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: MTB39 (TbH9FL)  
US-09-886-349A-14

Query Match 100.0%; Score 1949; DB 12; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.2e-145;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MYDFGALPEINARMYAGPGSASLVAAQAQMDWSVASDLFSAASAFQSVVWGLTVGSMIG	60
DB	1	MYDFGALPEINARMYAGPGSASLVAAQAQMDWSVASDLFSAASAFQSVVWGLTVGSMIG	60
QY	61	SSAGLMVAAASPYVAMSVTACQALTAQVAAAYETAYGLTVPPIVIAENRAELMI	120
DB	61	SSAGLMVAAASPYVAMSVTACQALTAQVAAAYETAYGLTVPPIVIAENRAELMI	120



Query Match  
100.0%; Score 1949; DB 14; Length 391;

```

Db      9  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 68
QY      61  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120
Db      69  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 128
QY      121  LIATNLLGQNTPATAVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 180
Db      129  LIATNLLGQNTPATAVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 188
QY      181  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 240
Db      189  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 248
QY      241  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAPAAAQVQTAQAQNGVRAMSSSLGSSG 300
Db      249  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAPAAAQVQTAQAQNGVRAMSSSLGSSG 308
QY      301  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 360
Db      309  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 368
QY      361  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 391
Db      369  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 399

```

## RESULT 6

```

US-09-886-349A-20
; Sequence 20, Application US/09886349A
; Publication No. US20040086523A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886,349A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein TbH9-Ra35 (designated MTB59P)
US-09-886-349A-20

```

```

Query Match      100.0%; Score 1949; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60
Db      9  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 68
QY      61  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120
Db      69  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 128
QY      121  LIATNLLGQNTPATAVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 180
Db      129  LIATNLLGQNTPATAVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 188
QY      181  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 240

```

```

Db      189  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 248
QY      241  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAPAAAQVQTAQAQNGVRAMSSSLGSSG 300
Db      249  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAPAAAQVQTAQAQNGVRAMSSSLGSSG 308
QY      301  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 360
Db      309  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 368
QY      361  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 391
Db      369  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 399

```

## RESULT 7

```

US-10-359-460-26
; Sequence 26, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
US-10-359-460-26

```

```

Query Match      100.0%; Score 1949; DB 14; Length 596;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60
Db      9  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 68
QY      61  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120
Db      69  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 128
QY      121  LIATNLLGQNTPATAVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 180
Db      129  LIATNLLGQNTPATAVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 188
QY      181  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 240
Db      189  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 248

```





```

; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, David C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 06/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; US-10-359-460-22

```

```

Query Match      100.0%; Score 1949; DB 14; Length 600;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 60
Db 9 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 68
QY 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
Db 69 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 128
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
Db 129 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 188
QY 181 LLEQAAAVEASDTPAAANQNMNNVPQALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 240
Db 189 LLEQAAAVEASDTPAAANQNMNNVPQALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 248
QY 241 MYSMANNHSMNTNSGVSMNTNTLSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 300
Db 249 MYSMANNHSMNTNSGVSMNTNTLSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 308
QY 301 LGGGVAANLGRAASVGSLSVPQAWAANAQAVTPAARALPLTSLTSAERGPQMLGGLPV 360
Db 309 LGGGVAANLGRAASVGSLSVPQAWAANAQAVTPAARALPLTSLTSAERGPQMLGGLPV 368
QY 361 GQMGARAGGGLSVLRVPPRPVYVPHSPAAAG 391
Db 369 GQMGARAGGGLSVLRVPPRPVYVPHSPAAAG 399

```

```

RESULT 11
US-10-369-983-2
; Sequence 2, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-008081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mutated
; OTHER INFORMATION: MTB32-MTB39F fusion protein (MTB32MutSA)
; US-10-369-983-2

```

```

Query Match      100.0%; Score 1949; DB 15; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.6e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 60
Db 333 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 392
QY 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
Db 393 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 452
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
Db 453 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 512
QY 181 LLEQAAAVEASDTPAAANQNMNNVPQALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 240
Db 513 LLEQAAAVEASDTPAAANQNMNNVPQALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 572
QY 241 MYSMANNHSMNTNSGVSMNTNTLSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 300
Db 573 MYSMANNHSMNTNSGVSMNTNTLSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 632
QY 301 LGGGVAANLGRAASVGSLSVPQAWAANAQAVTPAARALPLTSLTSAERGPQMLGGLPV 360
Db 633 LGGGVAANLGRAASVGSLSVPQAWAANAQAVTPAARALPLTSLTSAERGPQMLGGLPV 692
QY 361 GQMGARAGGGLSVLRVPPRPVYVPHSPAAAG 391
Db 693 GQMGARAGGGLSVLRVPPRPVYVPHSPAAAG 723

```

```

RESULT 12
US-09-886-349A-18
; Sequence 18, Application US/09886349A
; Publication No. US20040086523A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886,349A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1

```

```
; SEQ ID NO 18
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:NTB72MutSA
; OTHER INFORMATION: (Ral2-TbHp-Ra35MutSA)
US-09-886-349A-18

Query Match          100.0%; Score 1949; DB 12; Length 729;
Best Local Similarity 100.0%; Pred.No.2.6e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1   MYDFGLPPEIN SARMYAGPSASIVAAAQWMDVSADLFSASAQSVVWGLTVGSWG 60
Db      142 MYDFGLPPEIN SARNYAGPSASLVAAAQWMDVSADLFSASAQSVVWGLTVGSWG 201

Qy      61   SSAGLMVAASPYVAWMVTGAQAE LTA AQVRVAAAAAYETAYGLTVPVPPIENRAELM 120
Db      202 SSAGLMVAASPYVAWMVTGAQAE LTA AQVRVAAAAAYETAYGLTVPVPPIENRAELM 261

Qy      121 LIATNLILGONTPAIAI VNEAEYGEWQAODAAAMFCGYAAATATATATLLPPEAPEMTSAGS 180
Db      262 LIATNLILGONTPAIAI VNEAEYGEWQAODAAAMFCGYAAATATATATLLPPEAPEMTSAGS 321

Qy      181 LLEQAAAVEEASDTAAANQLNMNVPQALQQCLAQPTQTGTPTSSKLGLWKTVSPHRSPIN 240
Db      322 LLEQAAAVEEASDTAAANQLNMNVPQALQQCLAQPTQTGTPTSSKLGLWKTVSPHRSPIN 381

Qy      241 MYSMANNHMSMTNSGVSM TNLTLSSMLKGFAPAAAQAQVATAAQNGVPMSSLGSLGSSG 300
Db      382 MYSMANNHMSMTNSGVSM TNLTLSSMLKGFAPAAAQAQVATAAQNGVPMSSLGSLGSSG 441

Qy      301 LGGGVAAANI GRAASVGSLSVPOAWAAANQAVTPAARALPLTSLTSA AERPGQMGLGPV 360
Db      442 LGGGVAAANI GRAASVGSLSVPOAWAAANQAVTPAARALPLTSLTSA AERPGQMGLGPV 501

Qy      361 GQGAPAGGGLSGVL RVPRPVYMHSPAAG 391
Db      502 GQGAPAGGGLSGVL RVPRPVYMHSPAAG 532
```

```

RESULT 13
US-10-098-732A-18
; Sequence 18, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72FMutSA
; OTHER INFORMATION: (Ra12-TbH9-Ra35MutSA)
US-10-098-732A-18

```

Query Match	100.0%	Score 1349;	DB 14;	Length 729;
Best Local Similarity	100.0%	Pred. No. 2.6e-145;		
Matches 391;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps 0;				

Qy	1	MVDFGALPPEINSGARMYAGP	SGSASLVAAACQWDSVASD	LSAASAFOSVVMGLTVGSWIG	60
Db	142	MVDFGALPPEINSGARMYAGP	SGSASLVAAACQWDSVASD	LSAASAFOSVVMGLTVGSWIG	201
Qy	61	SSAGLVMVAASPYVMWSVT	AGAEATAAQVRVAAAAAYET	AYGLTVPPVIAENRAELMI	120
Db	202	SSAGLVMVAASPYVMWSVT	AGAEATAAQVRVAAAAAYET	AYGLTVPPVIAENRAELMI	261
Qy	121	LIATNLIGONTPAIVN	EABYGENWAQDAAMFGYAA	ATAATATATATLPPFEA	PEMTSAGG 180
Db	262	LIATNLIGONTPAIVN	EABYGENWAQDAAMFGYAA	ATAATATATLPPFEA	PEMTSAGG 321
Qy	181	LLEQAAAVEASDTAAAN	OLMNNVPQALQOLAQPTQ	GTTSSKLGGLWKTVP	HRSPISN 240
Db	322	LLEQAAAVEASDTAAAN	OLMNNVPQALQOLAQPTQ	GTTSSKLGGLWKTVP	HRSPISN 381
Qy	241	MVSMANNHMTNSGV	SWNTNLSMLKGFAPAAAA	QAVQTAAQNGVRAM	SLGSSLGSSG 300
Db	382	MVSMANNHMTNSGV	SWNTNLSMLKGFAPAAAA	QAVQTAAQNGVRAM	SLGSSLGSSG 441
Qy	301	LGGSVAANLGRAASV	SGLSVPQAWAAANQAVT	PAARALPLTSLTSA	AERGPQMLGLLPV 360
Db	442	LGGSVAANLGRAASV	SGLSVPQAWAAANQAVT	PAARALPLTSLTSA	AERGPQMLGLLPV 501
Qy	361	GQMGARAGGGLSGV	LRVPPRPVYVMPHSPA	AG 391	
Db	502	GQMGARAGGGLSGV	LRVPPRPVYVMPHSPA	AG 532	

```

RESULT 14
US-10-369-983-21
; Sequence 21, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeivy, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72F
US-10-369-983-21

```

Query Match	100.0%;	Score 1949;	DB 15;	Length 729;
Best Local Similarity	100.0%;	Pred. No. 2.6e-145;		
Matches 391:	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;

Qy	1	MVDFGALPPEINSAARMYAGPGSASLVAIAAQMWDSVASDLFSAASAFQSVVWGLTTVGWSMG	60
Db	142	MVDFGALPPEINSAARMYAGPGSASLVAIAAQMWDSVASDLFSAASAFQSVVWGLTTVGWSMG	201
Qy	61	SSAGLMVAASAPPYVAAWMSVTAGQAEELTAQQRVRAAAAYETAYGLTVPPVIAENRAELMI	120
Db	202	SSAGLMVAASAPPYVAAWMSVTAGQAEELTAQQRVRAAAAYETAYGLTVPPVIAENRAELMI	261
Qy	121	LIATNLLGONTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG	180
Db	262	LIATNLLGONTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG	321
Qy	181	LLEQAAAVEEASDTAAANQLMNNVPALQOLQAQPTQGTTPSSKLGGLMKTVPSPHSPISIN	240
Db	322	LLEQAAAVEEASDTAAANQLMNNVPALQOLQAQPTQGTTPSSKLGGLMKTVPSPHSPISIN	381



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:57:18 ; Search time 128 Seconds  
(without alignments)  
863.094 Million cell updates/sec

Title: US-09-724-685-107  
Perfect score: 1949  
Sequence: 1 MVDPCALPEINSARMYAGP.....SGVLRVPRPYMPHSPAAG 391.

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Capext 0.5

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq.29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1949	100.0	391	2	AAW32381	Myocobacte
2	1949	100.0	391	2	AAW32449	Myocobacte
3	1949	100.0	391	2	AAW43335	Myocobacte
4	1949	100.0	391	2	AAW81702	M. tuberc
5	1949	100.0	391	2	AAW04778	Myocobacte
6	1949	100.0	391	2	AAW38989	M. tuberc
7	1949	100.0	391	2	AAW39132	M. tuberc
8	1949	100.0	391	4	AAU01888	M. tuberc
9	1949	100.0	391	5	AAE29707	Myocobacte
10	1949	100.0	391	5	AAE17571	Myocobacte
11	1949	100.0	596	2	AAW32070	Myocobacte
12	1949	100.0	596	5	AAE29710	Myocobacte
13	1949	100.0	596	5	AAE17574	Myocobacte
14	1949	100.0	599	5	AAU74599	Antigenic
15	1949	100.0	600	2	AAW32068	Myocobacte
16	1949	100.0	600	5	AAU74597	Antigenic
17	1949	100.0	723	7	ADA26354	Myocobacte
18	1949	100.0	729	4	AAO22142	Ra12-H9-3
19	1949	100.0	729	5	AAE29709	Myocobacte
20	1949	100.0	729	5	AAE17573	Myocobacte
21	1949	100.0	729	7	ADA26374	Myocobacte
22	1949	100.0	744	4	AAU01902	M. tuberc
23	1949	100.0	813	7	ADA26367	Myocobacte
24	1949	100.0	815	4	AAU01904	M. tuberc
25	1949	100.0	825	7	ADA26366	Myocobacte

26	1949	100.0	875	7	ADA26365	Myocobacte
27	1949	100.0	930	5	AAE29731	Myocobacte
28	1949	100.0	930	7	ADA26364	Myocobacte
29	1949	100.0	1010	7	ADA26356	Myocobacte
30	1949	100.0	1016	7	ADA26370	M. bovis
31	1949	100.0	1022	7	ADA26369	Myocobacte
32	1949	100.0	1154	7	ADA26368	Myocobacte
33	1945	99.8	788	4	AAU01903	Myocobacte
34	1944	99.7	394	2	AAU04779	M. tuberc
35	1944	99.7	729	5	AAE29708	Myocobacte
36	1944	99.7	729	5	AAE17572	Myocobacte
37	1939	99.5	729	7	ADA26373	Myocobacte
38	1931	99.1	729	2	AAW32059	Myocobacte
39	1902.5	97.6	726	5	AAU74588	Antigenic
40	1652.5	84.8	396	2	AAW64337	Myocobacte
41	1652.5	84.8	396	2	AAW81704	M. tuberc
42	1652.5	84.8	396	2	AAW38991	M. tuberc
43	1652.5	84.8	396	2	AAW39134	M. tuberc
44	1583	81.2	393	6	ABU36968	Protein e
45	1583	81.2	393	6	ABU34531	Protein e

ALIGNMENTS

RESULT 1  
AAW32381  
ID AAW32381 standard; protein; 391 AA.  
XX  
AC AAW32381;  
XX  
DT 13-JAN-1998 (first entry)  
XX  
DE Mycobacterium tuberculosis antigen TbH-9FL.  
XX  
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;  
KW skin testing; M.tuberculosis.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO9709429-A2.  
XX  
PD 13-MAR-1997.  
XX  
PF 30-AUG-1996; 96WO-US014675.  
XX  
PR 01-SEP-1995; 95US-00523435.  
PR 22-SEP-1995; 95US-00532136.  
PR 22-MAR-1996; 96US-00620280.  
PR 05-JUN-1996; 96US-00658800.  
PR 12-JUL-1996; 96US-00680573.  
XX (CORI-) CORIXA CORP.  
XX  
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;  
PI Vedvick TH, Twardzik DR;  
XX  
DR WPI: 1937-192904/17.  
DR N-PSDB; AAT91455.  
XX  
CC New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -  
CC useful for diagnosis of M. tuberculosis infection.  
XX  
PS Example 3; Page 150-152; 190pp; English.  
XX  
CC A new immunogenic polypeptide has been developed comprising an  
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its  
CC variant differing only in conservative substitutions and/or  
CC modifications). The present sequence represents a M.tuberculosis antigen,  
CC TbH-9FL. The immunogenic polypeptide can be used to diagnose  
CC M.tuberculosis infection by forming complexes with specific antibodies in  
CC the sample. Fragments of DNA encoding the immunogenic polypeptide can be  
CC used as diagnostic primers or probes and agents that bind to the antigen,  
CC

CC especially monoclonal antibodies or equivalent polyclonal antibodies, are  
 CC also used for diagnosis

XX Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
 DB 1 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60

QY 61 SSAGLWVAASPYVWMSVTAGQELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
 DB 61 SSAGLWVAASPYVWMSVTAGQELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180  
 DB 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQQAQTOGTTPSSKLGGLWKTVPSPHRSPI 240  
 DB 181 LLEQAAVEEASDTAAANQLMNNVPOALQQAQTOGTTPSSKLGGLWKTVPSPHRSPI 240

QY 241 MVSMAHHMNTSGVSMNTLSSMLKGFAPAAAQVQTAQNGVRAMSSLGSSLGSSG 300  
 DB 241 MVSMAHHMNTSGVSMNTLSSMLKGFAPAAAQVQTAQNGVRAMSSLGSSLGSSG 300

QY 301 LGGGVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
 DB 301 LGGGVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391  
 DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

# RESULT 2

AAW32449  
 ID AAW32449 standard; protein; 391 AA.

AC AAW32449;

XX 09-JAN-1998 (first entry)

XX Mycobacterium tuberculosis antigen TbH-9FL.

XX Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;

XX skin testing; M.tuberculosis.

XX Mycobacterium tuberculosis.

XX WO9709428-A2.

XX 13-MAR-1997.

XX 30-AUG-1996; 96WO-US014674.

XX 01-SEP-1995; 95US-00523436.

XX 22-SEP-1995; 95US-00533634.

XX 22-MAR-1996; 96US-00620874.

XX 05-JUN-1996; 96US-00659683.

XX 12-JUL-1996; 96US-00680574.

XX (CORI-) CORIXA CORP.

XX Reed SG, Skeiky VA, Dillon DC, Campos-Neto A, Houghton R;

XX Vedvick TH, Twardzik DR;

XX WPI; 1997-192903/17.

XX N-PSDB; AAT91521.

PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are  
 PT useful in vaccines for prevention or treatment of tuberculosis, also for  
 PT diagnosis.

XX Example 3; Page 138-139; 168pp; English.

XX A new immunogenic polypeptide has been developed comprising an  
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its  
 CC variant differing only in conservative substitutions and/or  
 CC modifications). The present sequence represents a M.tuberculosis antigen,  
 CC TbH-9FL. The immunogenic protein, and fusion proteins containing one or  
 CC more of the proteins or one of the proteins plus ESAN-6, are useful in  
 CC vaccines, preferably when formulated with a non-specific adjuvant, to  
 CC induce an immune response against M.tuberculosis (for treatment or  
 CC prevention)

XX Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;

Best Local Similarity 100.0%; Pred. No. 1.2e-142;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60

DB 1 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60

QY 61 SSAGLWVAASPYVWMSVTAGQELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

DB 61 SSAGLWVAASPYVWMSVTAGQELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATLPPFEAPEMTSAGG 180

DB 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATLPPFEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQQAQTOGTTPSSKLGGLWKTVPSPHRSPI 240

DB 181 LLEQAAVEEASDTAAANQLMNNVPOALQQAQTOGTTPSSKLGGLWKTVPSPHRSPI 240

QY 241 MVSMAHHMNTSGVSMNTLSSMLKGFAPAAAQVQTAQNGVRAMSSLGSSLGSSG 300

DB 241 MVSMAHHMNTSGVSMNTLSSMLKGFAPAAAQVQTAQNGVRAMSSLGSSLGSSG 300

QY 301 LGGGVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360

DB 301 LGGGVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

# RESULT 3

AAW64335

ID AAW64335 standard; protein; 391 AA.

XX AAW64335;

XX 17-OCT-2003 (revised)

XX 09-NOV-1998 (first entry)

XX Mycobacterium tuberculosis antigen TbH-9FL.

XX Tuberculosis; infection; diagnosis; antigen; TbH-9FL.

XX Mycobacterium tuberculosis; strain H37Rv.

XX WO9816645-A2.

XX 23-APR-1998.

XX 07-OCT-1997; 97WO-US018214.

XX 11-OCT-1996; 96US-00729622.

```
PR 13-MAR-1997; 97US-00818111.
XX (CORI-) CORIXA CORP.
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX Vedvick TS, Twardzik DR, Lodes MJ;
XX WPI; 1998-251292/22.
XX N-PSDB; AAV44395.
XX
XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used to
XX develop products for the detection of M. tuberculosis infection and
XX diagnosis of tuberculosis.
XX
XX Example 3; Page 133-135; 250pp; English.
XX
XX This polypeptide comprises Mycobacterium tuberculosis antigen TBH-9FL. It
XX is encoded by genomic DNA (see AAV44395) isolated from a M. tuberculosis
XX strain H37Rv genomic library using a probe from clone TBH-9 (see
XX AAV44371). The invention relates to compositions and methods for
XX diagnosing tuberculosis. It provides polypeptides (see AAV44291-W64379)
XX comprising an antigenic portion of a soluble M. tuberculosis antigen, or
XX an immunogenic portion of an M. tuberculosis antigen, as well as DNA
XX sequences encoding such polypeptides, recombinant expression vectors and
XX transformed or transfected host cells. Also claimed are methods and
XX diagnostic kits for detecting host M. tuberculosis infection in a patient
XX using these polypeptides, antibodies or oligonucleotide probes and
XX primers, for the diagnosis of tuberculosis. (Updated on 17-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 391 AA;
XX
XX Query Match 100.0%; Score 1949; DB 2; Length 391;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-142;
XX Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWIG 60
DB 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWIG 60
QY 61 SSAGLMVAASPYVAMSVTAGQELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
DB 61 SSAGLMVAASPYVAMSVTAGQELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPPEAPEMTSAGG 180
DB 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPPEAPEMTSAGG 180
QY 181 LLEQAAVEEASD7TAAANQNMNVPAQLQQLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
DB 181 LLEQAAVEEASD7TAAANQNMNVPAQLQQLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
QY 241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSSG 300
DB 241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSSG 300
QY 301 LGGVVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360
DB 301 LGGVVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360
QY 361 GQMGARAGGGLSGVLRVPPRPVPHSPAAAG 391
DB 361 GQMGARAGGGLSGVLRVPPRPVPHSPAAAG 391
XX
XX RESULT 4
XX AAW81702
XX ID AAW81702 standard; protein; 391 AA.
XX AC AAW81702;
XX XX
XX DT 27-JAN-1999 (first entry)
XX XX
```

```
DE M. tuberculosis immunogenic polypeptide TBH-9FL.
XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
XX vaccine; pharmaceutical; infection; diagnosis.
XX
XX Mycobacterium tuberculosis.
XX OS
XX WO9816646-A2.
XX PN
XX 23-APR-1998.
XX PD
XX 07-OCT-1997; 97WO-US018293.
XX PF
XX 11-OCT-1996; 96US-00730510.
XX PR
XX 13-MAR-1997; 97US-00818112.
XX PS
XX (CORI-) CORIXA CORP.
XX PA
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX PI Vedvick TS, Twardzik DR, Lodes MJ;
XX PSDB; AAV64503.
XX WPI; 1998-261042/23.
XX N-PSDB; AAV64503.
XX
XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
XX develop products for the detection of M. tuberculosis infection and for
XX diagnosis, treatment and prevention of tuberculosis.
XX
XX Example 3B; Page 128-129; 230pp; English.
XX
XX This sequence represents an immunogenic portion of a soluble
XX Mycobacterium tuberculosis (MT) antigen which can be used in a method for
XX inducing protective immunity against tuberculosis (TB). This sequence can
XX be formulated into vaccines and/or pharmaceutical compositions for
XX immunising against M. tuberculosis infection or may be used for the
XX diagnosis of tuberculosis
XX
XX Sequence 391 AA;
XX
XX Query Match 100.0%; Score 1949; DB 2; Length 391;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-142;
XX Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWIG 60
DB 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWIG 60
QY 61 SSAGLMVAASPYVAMSVTAGQELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
DB 61 SSAGLMVAASPYVAMSVTAGQELTAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
DB 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
QY 181 LLEQAAVEEASD7TAAANQNMNVPAQLQQLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
DB 181 LLEQAAVEEASD7TAAANQNMNVPAQLQQLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
QY 241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSSG 300
DB 241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSSG 300
QY 301 LGGVVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360
DB 301 LGGVVAANLGRAASVGSLSVPOAWAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360
QY 361 GQMGARAGGGLSGVLRVPPRPVPHSPAAAG 391
DB 361 GQMGARAGGGLSGVLRVPPRPVPHSPAAAG 391
XX
XX RESULT 5
```

AAV04778  
 ID AAY04778 standard; protein; 391 AA.  
 XX  
 AC AAY04778;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE Mycobacterium species protein sequence 5R.  
 XX  
 KW Secreted protein; Mycobacterium; primer: PCR; amplification; probe;  
 KW hybridisation; detection; vaccine; immunisation; infection.  
 XX  
 OS Mycobacterium sp.  
 XX  
 PN WO9909186-A2.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 14-AUG-1998; 98WO-FR001813.  
 XX  
 PR 14-AUG-1997; 97FR-00010404.  
 PR 11-SEP-1997; 97FR-00011325.  
 XX  
 PA (INSP) INST PASTEUR.  
 XX  
 XX Gicquel B, Portnoie D, Lim E, Pelicic V, Guigueno A;  
 PI Goguet de La Salmoniere Y;  
 XX  
 DR WPI; 1999-181045/15.  
 DR N-PSDB; AAX34030.  
 XX  
 PT Mycobacterial DNA vectors containing reporter constructs - for  
 PT identifying coding or promoter sequences involved in infection-associated  
 PT protein expression.  
 XX  
 PS Claim 32; Fig 5R; 309pp; French.  
 XX  
 CC Sequences AAY04742-Y05000 and AAY07201-Y07204 represent secreted proteins  
 CC from various Mycobacterium species microorganisms. The encoding  
 CC nucleotide sequences can be used as primers and probes for methods for  
 CC detecting and identifying mycobacteria, especially belonging to the M.  
 CC tuberculosis complex. The encoded proteins can be used in vaccines for  
 CC immunisation against a bacterial or viral infection  
 XX  
 SQ Sequence 391 AA;  
 Query Match 100.0%; Score 1949; DB 2; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60  
 Db 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60  
 QY 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
 Db 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
 QY 121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATATATATATLLPPEAPEMTSAGG 180  
 Db 121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATATATATATLLPPEAPEMTSAGG 180  
 QY 181 LLEQAAVEASDSTAANQNMNVPOALQLOAQTOGTTTSSKLGGLWKTVPSPHSPISN 240  
 Db 181 LLEQAAVEASDSTAANQNMNVPOALQLOAQTOGTTTSSKLGGLWKTVPSPHSPISN 240  
 QY 241 MVSMANNHMTSGVSMNTLTSSMLKGFAPAAAAQAVTAAQNGVRAMSSLGSSLGSSG 300  
 Db 241 MVSMANNHMTSGVSMNTLTSSMLKGFAPAAAAQAVTAAQNGVRAMSSLGSSLGSSG 300  
 QY 301 LGGVAAANTCGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAABRGPCQMLGGLPV 360  
 Db 301 LGGVAAANTCGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAABRGPCQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLAVPPRPYPVMPHSPAAG 391  
 Db 361 GQMGARAGGGLSGVLAVPPRPYPVMPHSPAAG 391  
 RESULT 6  
 AAY38989  
 ID AAY38989 standard; protein; 391 AA.  
 XX  
 AC AAY38989;  
 XX  
 DT 05-NOV-1999 (first entry)  
 XX  
 DE M. tuberculosis recombinant antigen protein TbH-9FL.  
 XX  
 KW Antigen; diagnosis; detection; infection; antibody; immunisation;  
 KW vaccine; immunity.  
 XX  
 OS Mycobacterium tuberculosis.  
 XX  
 PN WO9942118-A2.  
 XX  
 PD 26-AUG-1999.  
 XX  
 PF 17-FEB-1999; 99WO-US003265.  
 PR 18-FEB-1998; 98US-00024753.  
 PR 05-MAY-1998; 98US-00072596.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedwick TS, Twardzik DR, Lodes MJ, Hendrickson RC;  
 XX  
 DR WPI; 1999-527416/44.  
 DR N-PSDB; AAZ19093.  
 XX  
 PT New polypeptide comprising antigenic portions of M. tuberculosis.  
 PS Example 3; Page 168-169; 323pp; English.  
 XX  
 CC This invention describes novel recombinant antigens and their encoding  
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel  
 CC polypeptides are useful for detecting M. tuberculosis infection in a  
 CC biological sample by detecting antibodies which bind with the  
 CC polypeptides, and are useful as vaccines for immunizing against M.  
 CC tuberculosis infection. The new detection methods are needed as current  
 CC vaccination strategies do not provide 100% immunity  
 XX  
 SQ Sequence 391 AA;  
 Query Match 100.0%; Score 1949; DB 2; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60  
 Db 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60  
 QY 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
 Db 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
 QY 121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180  
 Db 121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180  
 QY 181 LLEQAAVEASDSTAANQNMNVPOALQLOAQTOGTTTSSKLGGLWKTVPSPHSPISN 240  
 Db 181 LLEQAAVEASDSTAANQNMNVPOALQLOAQTOGTTTSSKLGGLWKTVPSPHSPISN 240  
 QY 241 MVSMANNHMTSGVSMNTLTSSMLKGFAPAAAAQAVTAAQNGVRAMSSLGSSLGSSG 300



```

Db      241  MYSMANNHSMNTNSGVSMNTLSSMLKGFAPAAAAQAVOTTAQNGVRAMSSLGSSGSSG 300
QY      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPGQMLGGLPV 360
Db      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPGQMLGGLPV 360
QY      361  GQMGARAGGGLSGVLRVPPRYVMPHSPAAG 391
Db      361  GQMGARAGGGLSGVLRVPPRYVMPHSPAAG 391

RESULT 7
AAU01888
ID      AAY39132 standard; protein; 391 AA.
XX
AC      AAY39132;
XX
DT      05-NOV-1999 (first entry)
XX
DE      M. tuberculosis antigen Tbh-9FL amino acid sequence.
XX
KW      Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW      immunotherapy; diagnosis; immunisation; vaccine; infection;
KW      immune response; skin test.
XX
OS      Mycobacterium tuberculosis.
XX
PN      WO9942076-A2.
XX
PD      26-AUG-1999.
XX
PF      17-FEB-1999; 99WO-US003268.
XX
PR      18-FEB-1998; 98US-00025197.
PR      05-MAY-1998; 98US-00072967.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI      Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
DR      WPI; 1999-527409/44.
DR      N-PSDB; AAZ19305.
XX
PT      New antigens from Mycobacterium tuberculosis useful in diagnostic skin
PT      tests and protective or therapeutic vaccines or compositions.
XX
PS      Example 3; Page 123-124; 299pp; English.
XX
CC      The present invention describes polypeptides comprising an immunogenic
CC      part of a Mycobacterium tuberculosis antigen (Ag). Also described are
CC      vaccines and fusion protein containing M. tuberculosis Ag's. M.
CC      tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
CC      polypeptides fragments, can be used in pharmaceutical compositions or
CC      vaccines to generate a protective or therapeutic immune response to M.
CC      tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
CC      Ag can induce proliferation of, or cytokine secretion by, T, B or natural
CC      killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249
CC      to AAZ19460 and AAY39083 to AAY39225 are used in the exemplification of
CC      the present invention
XX
SQ      Sequence 391 AA;
XX
Query Match      100.0%; Score 1949; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.2e-142;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1  MVDFGALPPEINARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSGWIG 60
Db      1  MVDFGALPPEINARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSGWIG 60
QY      61  SSAGLWVAASPYYVAMSVTAGQAELETAQVRVAAAAAYETAYGLTVPPIAENRAELMI 120

```

```

Db      61  SSAGLWVAASPYYVAMSVTAGQAELETAQVRVAAAAAYETAYGLTVPPIAENRAELMI 120
QY      121  LIATNLGQNTPALAVNEAEYGEWAAQDAAMFGYAAATATATATATLLPPEAEPMTSAGG 180
Db      121  LIATNLGQNTPALAVNEAEYGEWAAQDAAMFGYAAATATATATATLLPPEAEPMTSAGG 180
QY      181  LLEQAAVEEASDTAAANQLMNNVPQALQLOAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240
Db      181  LLEQAAVEEASDTAAANQLMNNVPQALQLOAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240
QY      241  MVSMAHHMNTNSGVSMNTLSSMLKGFAPAAAAQAVOTTAQNGVRAMSSLGSSGSSG 300
Db      241  MVSMAHHMNTNSGVSMNTLSSMLKGFAPAAAAQAVOTTAQNGVRAMSSLGSSGSSG 300
QY      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPGQMLGGLPV 360
Db      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPGQMLGGLPV 360
QY      361  GQMGARAGGGLSGVLRVPPRYVMPHSPAAG 391
Db      361  GQMGARAGGGLSGVLRVPPRYVMPHSPAAG 391

RESULT 8
AAU01888
ID      AAU01888 standard; protein; 391 AA.
XX
AC      AAU01888;
XX
DT      29-AUG-2001 (first entry)
XX
DE      M. tuberculosis antigen Tbh9 (Mtb39A).
XX
KW      Tbh9; Mtb39A; antigen; vaccine; tuberculosis; AIDS;
KW      acquired immunodeficiency disease.
XX
OS      Mycobacterium tuberculosis.
XX
PN      WO200124820-A1.
XX
PD      12-APR-2001.
XX
PF      10-OCT-2000; 2000WO-US028095.
XX
PR      07-OCT-1999; 99US-0158338P.
PR      07-OCT-1999; 99US-0158425P.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Skeiky Y, Reed S, Houghton RL, Mcneill PD, Dillon DC, Lodes ML;
XX
DR      WPI; 2001-290576/30.
DR      N-PSDB; AAS03779.
XX
PT      Vaccinating against Mycobacteria infections in mammals using fusion
PT      proteins comprising combinations of heterologous antigens.
XX
PS      Example 2; Page 151-152; 169pp; English.
XX
CC      The sequence represents Mycobacterium tuberculosis Tbh9 (also known as
CC      Mtb39A), an M. tuberculosis antigen. Compositions comprising at least 2
CC      heterologous antigens, as a fusion protein, and vectors expressing the
CC      fusion proteins are used as vaccines to prophylactically immunise mammals
CC      (especially humans) against infection by Mycobacteria. The compositions
CC      contain at least 2 heterologous antigens that increase the serological
CC      sensitivity of individuals infected with tuberculosis, a disease
CC      frequently affecting patients with acquired immunodeficiency disease,
CC      AIDS
XX
SQ      Sequence 391 AA;
XX
Query Match      100.0%; Score 1949; DB 4; Length 391;

```

Best Local Similarity 100.0%; Pred. No. 1.2e-142; Mismatches 0; Indels 0; Gaps 0;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60  
DB 1 MVDFGALPPEINSARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60

QY 61 SSAGLMVAASAPFYVAMSVTAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLMVAASAPFYVAMSVTAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180  
DB 121 LIATNLLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180

QY 181 LLEQAAAVEEASDTAAANQNMNVFQALQQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240  
DB 181 LLEQAAAVEEASDTAAANQNMNVFQALQQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240

QY 241 MVSMAHNMSTNGSVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSSLSGSSG 300  
DB 241 MVSMAHNMSTNGSVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSSLSGSSG 300

QY 301 LGGVVAANLGRAASVGSLSVPOAWAANOAVTPAARALPLTSLTSAERCPGQMLGGLPV 360  
DB 301 LGGVVAANLGRAASVGSLSVPOAWAANOAVTPAARALPLTSLTSAERCPGQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391

RESULT 9  
AAE29707  
ID AAE29707 standard; protein; 391 AA.  
XX AC AAE29707;  
XX DT 27-JAN-2003 (first entry)  
XX DE Mycobacterium sp. TbH9FL antigenic protein.  
XX KW Vaccine; immunity; diagnostic agent; gene therapy; TbH9FL antigen.  
XX OS Mycobacterium sp.  
XX PN WO200272792-A2.  
XX PD 19-SEP-2002.  
XX PF 13-MAR-2002; 2002WO-US008223.  
XX PR 13-MAR-2001; 2001US-0275837P.  
XX PA (CORI-) CORIXA CORP.  
XX PI Skeiky Y, Brannon M, Guderian J;  
XX DR WPI; 2002-759844/82.  
XX DR N-PSDB; AAD47082.  
XX FT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeifP,  
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity  
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium  
PT tuberculosis.  
XX PS Disclosure; Page 86-87; 155pp; English.  
XX CC The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
CC polypeptide or its fragment. The Leishmania polynucleotide is selected

CC from TSA, LeifP, M15, and 6H polynucleotides. Sequences of the invention  
CC are used in methods for eliciting immune response in mammals. They are  
CC useful as vaccines to elicit protective immunity against pathogenic  
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
CC polypeptides are used for enhancing the expression of polynucleotides, as  
CC in vivo diagnostic agents and for raising antibodies in a non-human  
CC animal. The invention is used in gene therapy. The present sequence is  
CC Mycobacterium sp. TbH9FL antigenic protein  
XX  
SQ Sequence 391 AA;

Query Watch 100.0%; Score 1949; DB 5; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.2e-142;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60  
DB 1 MVDFGALPPEINSARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60

QY 61 SSAGLMVAASAPFYVAMSVTAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLMVAASAPFYVAMSVTAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180  
DB 121 LIATNLLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180

QY 181 LLEQAAAVEEASDTAAANQNMNVFQALQQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240  
DB 181 LLEQAAAVEEASDTAAANQNMNVFQALQQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240

QY 241 MVSMAHNMSTNGSVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSSLSGSSG 300  
DB 241 MVSMAHNMSTNGSVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSSLSGSSG 300

QY 301 LGGVVAANLGRAASVGSLSVPOAWAANOAVTPAARALPLTSLTSAERCPGQMLGGLPV 360  
DB 301 LGGVVAANLGRAASVGSLSVPOAWAANOAVTPAARALPLTSLTSAERCPGQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391  
DB 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391

RESULT 10  
AAE17571  
ID AAE17571 standard; protein; 391 AA.  
XX AC AAE17571;  
XX DT 22-APR-2002 (first entry)  
XX DE Mycobacterium species MTB39 (TbH9) protein #2.  
XX KW Fusion protein; antigen; serological sensitivity; immune response;  
XX KW tuberculosis; infection; vaccine; MTB39; TbH9 protein.  
XX OS Mycobacterium sp.  
XX PN WO200198460-A2.  
XX PD 27-DEC-2001.  
XX PF 20-JUN-2001; 2001WO-US019959.  
XX PR 20-JUN-2000; 2000US-00597796.  
XX PR 01-FEB-2001; 2001US-0265737P.  
XX PA (CORI-) CORIXA CORP.  
XX PI Skeiky Y, Reed S, Alderson M;  
XX DR WPI; 2002-147798/19.

DR N-PSDB; AAD28341.  
 XX Composition comprising MTB39 antigen and MTB32A antigen from  
 PT Mycobacterium species, useful for eliciting immune response in a subject.  
 XX  
 PS Claim 83; Page 102-103; 136pp; English.  
 XX  
 CC The present invention relates to fusion proteins containing at least two  
 CC Mycobacterium species antigens, nucleotides encoding them and  
 CC compositions comprising such fusion proteins. The present invention  
 CC particularly relates to nucleic acids encoding fusion proteins that  
 CC include two or more individual M. tuberculosis antigens which increase  
 CC the serological sensitivity of sera from individuals infected with  
 CC tuberculosis and methods for their use in diagnosis, prevention and  
 CC treatment of tuberculosis infection. Sequences of the invention are  
 CC useful for eliciting an immune response in a mammal, e.g., human,  
 CC immunised with BCG. They are useful in the diagnosis, treatment and  
 CC prevention of Mycobacterium infection. The fusion proteins and the  
 CC polynucleotides are useful as diagnostic tools in patients infected with  
 CC Mycobacterium, in vitro and in vivo assays for detecting humoral  
 CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
 CC diagnosis of an infection or monitoring of disease progression, as  
 CC immunogens to generate or elicit a protective immune response in a  
 CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
 CC animal. Sequences of the invention are also used as vaccines. MTB32A  
 CC fusion proteins of the invention are useful as in vivo diagnostic agents  
 CC for intradermal skin test. The present sequence is Mycobacterium species  
 CC MTB39 (TbH9) protein  
 XX  
 SQ Sequence 391 AA;  
 Query Match 100.0%; Score 1949; DB 5; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MYDFGALPPPEINSAARMYAGPGSASLVAAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
 Db 1 MYDFGALPPPEINSAARMYAGPGSASLVAAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
 QY 61 SSAGLMVAASPYVAMSVTAGQAELTAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120  
 Db 61 SSAGLMVAASPYVAMSVTAGQAELTAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120  
 QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 180  
 Db 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 180  
 QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQOLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240  
 Db 181 LLEQAAVEEASDTAAANQLMNNVPOALQOLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240  
 QY 241 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 300  
 Db 241 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 300  
 QY 301 LGGVVAANLGRAASVGLSVPOQAAWAAQAVTPAARALPLTSLTSAARPGQWMLGGLPV 360  
 Db 301 LGGVVAANLGRAASVGLSVPOQAAWAAQAVTPAARALPLTSLTSAARPGQWMLGGLPV 360  
 QY 361 GQMGARAGGGLSVGLRVPPRPVMPHSPAAG 391  
 Db 361 GQMGARAGGGLSVGLRVPPRPVMPHSPAAG 391  
 RESULT 11  
 AAY32070  
 ID AAY32070 standard; protein; 596 AA.  
 XX  
 AC AAY32070;  
 XX  
 DT 17-JAN-2000 (first entry)  
 XX  
 DE Mycobacterium tuberculosis antigen fusion protein Mtb59f.

XX Tuberculosis; antigen; fusion protein; Mtb59f; TbH9; Ra35; diagnosis;  
 XX therapy; vaccine; immunogen.  
 XX Mycobacterium tuberculosis.  
 OS  
 XX Key Location/Qualifiers  
 FH Peptide 1..8  
 FT /note= "Met/His tag"  
 FT Protein 9..140  
 FT /note= "Ra12"  
 FT Protein 143..596  
 FT /note= "TbH9"  
 XX WO9951748-A2.  
 PN 14-OCT-1999.  
 PD 07-APR-1999; 99WO-US007717.  
 PF 07-APR-1998; 98US-00056556.  
 PR 30-DEC-1998; 98US-00223040.  
 XX (CORI-) CORIXA CORP.  
 PA Skeiky YAW, Alderson M, Campos-Neto A;  
 XX WPI; 1999-601610/51.  
 DR N-PSDB; AAZ20205.  
 XX New fusion proteins useful for diagnosis, prevention and treatment of  
 PT tuberculosis.  
 FT Claim 1; Fig 12A-B; 83pp; English.  
 XX  
 PS This sequence represents a recombinant Mycobacterium tuberculosis bi-  
 CC antigen fusion protein, termed Mtb59f, composed of the antigens TbH9 and  
 CC Ra35. The fusion protein is expressed in host cells using a vector  
 CC carrying a polynucleotide (see AAZ20205) comprising the coding sequences  
 CC for the 2 antigens. The invention provides fusion proteins (see AAY32059-  
 CC 71) containing at least 2 M. tuberculosis antigens. The new fusion  
 CC proteins and polynucleotides encoding them are useful as vaccines for  
 CC preventing tuberculosis (claimed), for diagnosis (via in vitro assays or  
 CC intradermal skin tests for detection of anti-M. tuberculosis antibodies),  
 CC monitoring of disease progression, and treatment of tuberculosis. They  
 CC are more effective immunogens than mixtures of the individual protein  
 CC components  
 XX  
 SQ Sequence 596 AA;  
 Query Match 100.0%; Score 1949; DB 2; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MYDFGALPPPEINSAARMYAGPGSASLVAAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
 Db 9 MYDFGALPPPEINSAARMYAGPGSASLVAAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 65  
 QY 61 SSAGLMVAASPYVAMSVTAGQAELTAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120  
 Db 69 SSAGLMVAASPYVAMSVTAGQAELTAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 128  
 QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 180  
 Db 129 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 188  
 QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQOLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240  
 Db 189 LLEQAAVEEASDTAAANQLMNNVPOALQOLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 248  
 QY 241 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 300  
 Db 249 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 308



CC immunised with BCG. They are useful in the diagnosis, treatment and  
 CC prevention of Mycobacterium infection. The fusion proteins and the  
 CC polynucleotides are useful as diagnostic tools in patients infected with  
 CC Mycobacterium, in vitro and in vivo assays for detecting humoral  
 CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
 CC diagnosis of an infection or monitoring of disease progression, as  
 CC immunogens to generate or elicit a protective immune response in a  
 CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
 CC animal. Sequences of the invention are also used as vaccines. MTB32A  
 CC fusion proteins of the invention are useful as in vivo diagnostic agents  
 CC for intradermal skin test. The present sequence is Mycobacterium species  
 CC MTB59f (TbH9-Ra35) fusion protein  
 XX  
 XX Sequence 596 AA;

Query Match 100.0%; Score 1949; DB 5; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINARMYAGPGSASLVAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
 Db 9 MVDGALPPEINARMYAGPGSASLVAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 68  
 QY 61 SSAGLMVAASPYVAVMSVTAQAEELTAQAQVRAAAAYETAYGLTVPPPIAENRAELMI 120  
 Db 69 SSAGLMVAASPYVAVMSVTAQAEELTAQAQVRAAAAYETAYGLTVPPPIAENRAELMI 128  
 QY 121 LIATNLGQNTPAIAVNEAEGEMWAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 180  
 Db 129 LIATNLGQNTPAIAVNEAEGEMWAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 188  
 QY 181 LLEQAAVEEASDTAAANQLMNVPAQALQAOPTQGTTPSSKLGKLTVPSPHSPISN 240  
 Db 189 LLEQAAVEEASDTAAANQLMNVPAQALQAOPTQGTTPSSKLGKLTVPSPHSPISN 248  
 QY 241 MVSMAHNMHMTNSGVSMNTLSMNLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 300  
 Db 249 MVSMAHNMHMTNSGVSMNTLSMNLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 308  
 QY 361 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 391  
 Db 369 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 399

RESULT 14  
 AAU74599

ID AAU74599 standard; protein; 599 AA.

XX AAU74599;

XX 29-AUG-2003 (revised)

DT 08-MAY-2002 (first entry)

XX Antigenic fusion protein Tb59-Ra35 (Mtb59f).

XX Fusion protein; tuberculosis; Mycobacterium tuberculosis;  
 XX tuberculostatic; immunogen; vaccine; Tb59-Ra35; Mtb59f.

XX Mycobacterium tuberculosis.  
 OS Chimeric.

XX Key Location/Qualifiers

XX Misc-difference 597

FT /label= OTHER

FT /note= "OTHER= Xaa. Xaa= In frame stop codon"

XX US2002009459-A1.

XX 24-JAN-2002.

RESULT 15

XX 07-APR-1999; 99US-00287849.  
 XX 13-MAR-1997; 97US-00818112.  
 PR 01-OCT-1997; 97US-00942578.  
 PR 18-FEB-1998; 98US-00025197.  
 PR 07-APR-1998; 98US-00056556.  
 PR 30-DEC-1998; 98US-00223040.  
 XX (REED//) REED S. G.  
 PA (SKEI//) SKEIKY Y. A.  
 PA (DILL//) DILLON D. C.  
 PA (ALDE//) ALDERSON M. A.  
 PA (CAMP//) CAMPOS-NETO A.  
 XX Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;  
 XX WPI; 2002-171134/22.  
 DR N-PSDB; ABK14139.  
 XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for  
 PT diagnosing, treating or preventing M. tuberculosis infection,  
 PT particularly as vaccine for treating or preventing tuberculosis.  
 XX Claim 1; Fig 12; 62pp; English.  
 XX The invention relates to a purified polypeptide which induces an immune  
 CC response of Mycobacterium tuberculosis. Polypeptides of the invention are  
 CC useful for diagnosing, treating or preventing M. tuberculosis infection,  
 CC particularly tuberculosis infection. In particular, the polypeptides are  
 CC useful as a vaccine formulation with an adjuvant to afford long-term  
 CC protection in animals against the development of tuberculosis. The  
 CC protein coding sequence may be used to encode a protein product for use  
 CC as an immunogen to induce and/or enhance an immune response to M.  
 CC tuberculosis. This sequence represents an M. tuberculosis fusion protein  
 CC of the invention. (Updated on 29-AUG-2003 to standardise OS field)  
 XX Sequence 599 AA;

Query Match 100.0%; Score 1949; DB 5; Length 599;

Best Local Similarity 100.0%; Pred. No. 2.1e-142;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINARMYAGPGSASLVAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
 Db 9 MVDGALPPEINARMYAGPGSASLVAQAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 68  
 QY 61 SSAGLMVAASPYVAVMSVTAQAEELTAQAQVRAAAAYETAYGLTVPPPIAENRAELMI 120  
 Db 69 SSAGLMVAASPYVAVMSVTAQAEELTAQAQVRAAAAYETAYGLTVPPPIAENRAELMI 128  
 QY 121 LIATNLGQNTPAIAVNEAEGEMWAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 180  
 Db 129 LIATNLGQNTPAIAVNEAEGEMWAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 188  
 QY 181 LLEQAAVEEASDTAAANQLMNVPAQALQAOPTQGTTPSSKLGKLTVPSPHSPISN 240  
 Db 189 LLEQAAVEEASDTAAANQLMNVPAQALQAOPTQGTTPSSKLGKLTVPSPHSPISN 248  
 QY 241 MVSMAHNMHMTNSGVSMNTLSMNLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 300  
 Db 249 MVSMAHNMHMTNSGVSMNTLSMNLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 308  
 QY 301 LGGGVAANLGRAASVGSLSVPAQAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360  
 Db 309 LGGGVAANLGRAASVGSLSVPAQAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 368  
 QY 361 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 391  
 Db 369 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 399



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:08:19 ; Search time 34 seconds  
(without alignments)  
593.699 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDGALPEINSARMYAGP.....SGVLRVPRPYVMHPSPAAG 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.\*
- 1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/prodata/2/iaa/PCUTUS\_COMB.pep.\*
  - 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1949	100.0	391	3	US-08-818-112-107
2	1949	100.0	391	4	US-08-818-111-102
3	1949	100.0	391	4	US-09-056-556-107
4	1949	100.0	391	4	US-09-072-596-102
5	1949	100.0	391	4	US-09-072-596-102
6	1949	100.0	596	4	US-09-287-849-26
7	1949	100.0	600	4	US-09-287-849-22
8	1944	99.7	729	4	US-09-223-040-2
9	1944	99.7	729	4	US-09-287-849-2
10	1652.5	84.8	396	3	US-08-818-112-111
11	1652.5	84.8	396	4	US-08-818-111-106
12	1652.5	84.8	396	4	US-09-056-556-111
13	1652.5	84.8	396	4	US-09-072-596-106
14	1652.5	84.8	396	4	US-09-072-596-111
15	1486.5	76.3	359	3	US-08-818-112-109
16	1486.5	76.3	359	4	US-08-818-111-104
17	1486.5	76.3	359	4	US-09-056-556-109
18	1486.5	76.3	359	4	US-09-072-596-104
19	1486.5	76.3	359	4	US-09-072-596-109
20	1187.5	60.9	358	4	US-09-287-849-8
21	1187	60.9	263	3	US-08-818-112-91
22	1187	60.9	263	4	US-08-818-111-92
23	1187	60.9	263	4	US-09-056-556-91
24	1187	60.9	263	4	US-09-072-596-92
25	1187	60.9	263	4	US-09-072-596-91
26	766.5	39.3	400	4	US-09-073-009-126
27	766.5	39.3	400	4	US-09-073-010-126

28	604	31.0	710	4	US-09-287-849-16
29	604	31.0	856	4	US-09-287-849-12
30	603	30.9	423	4	US-09-073-009-142
31	603	30.9	423	4	US-09-073-010-142
32	424.5	21.8	943	4	US-09-477-135A-131
33	424	21.8	141	4	US-09-073-009-15
34	424	21.8	141	4	US-09-073-010-15
35	381.5	19.6	204	4	US-08-311-731A-57
36	377.5	19.4	208	4	US-08-311-731A-208
37	317	16.3	371	4	US-09-050-739-92
38	314	16.1	368	3	US-08-818-112-114
39	314	16.1	368	4	US-08-818-111-109
40	314	16.1	368	4	US-09-056-556-114
41	314	16.1	368	4	US-09-072-596-109
42	314	16.1	368	4	US-09-072-967-114
43	261	13.4	1271	1	US-08-095-734-2
44	261	13.4	1271	2	US-08-444-623-2
45	261	13.4	1271	3	US-08-471-869-2

## ALIGNMENTS

RESULT 1  
US-08-818-112-107  
; Sequence 107, Application US/08818112  
; Patent No. 6290969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas R.  
; APPLICANT: Twardzik, Daniel S.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 822-4900  
; TELEFAX: (206) 822-6031  
; INFORMATION FOR SEQ ID NO: 107:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 391 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-818-112-107

Query Match 100.0%; Score 1949; DB 3; Length 391;  
Best Local Similarity 100.0%; Pred. No. 2.9e-154;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60  
DB 1 MVDGALPEINARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60  
QY 61 SSAGLWMAAASPYYAWMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
DB 61 SSAGLWMAAASPYYAWMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
QY 121 LIATNLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180  
DB 121 LIATNLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180  
QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240  
DB 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240  
QY 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300  
DB 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300  
QY 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPOMLGLPV 360  
DB 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPOMLGLPV 360  
QY 361 QMGARAGGGLSGVLVPPRPYPVMPHSPAAG 391  
DB 361 QMGARAGGGLSGVLVPPRPYPVMPHSPAAG 391

## RESULT 2

US-08-818-111-102  
; Sequence 102, Application US/08818111  
; Patent No. 6338852

## GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedwick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,111  
FILING DATE: 13-MAR-1997

## CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C6  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 102:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-818-111-102

Query Match 100.0%; Score 1949; DB 4; Length 391;  
Best Local Similarity 100.0%; Pred. No. 2.9e-154;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60  
DB 1 MVDGALPEINARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60  
QY 61 SSAGLWMAAASPYYAWMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
DB 61 SSAGLWMAAASPYYAWMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120  
QY 121 LIATNLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATLPPFEAPEMTSAGG 180  
DB 121 LIATNLGONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATLPPFEAPEMTSAGG 180  
QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240  
DB 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240  
QY 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300  
DB 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300  
QY 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPOMLGLPV 360  
DB 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPOMLGLPV 360  
QY 361 QMGARAGGGLSGVLVPPRPYPVMPHSPAAG 391  
DB 361 QMGARAGGGLSGVLVPPRPYPVMPHSPAAG 391

## RESULT 3

US-09-056-556-107  
; Sequence 107, Application US/09056556  
; Patent No. 6350456

## GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-APR-1998

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

TREATME



US-09-072-967-107  
; Sequence 107, Application US/09072967  
; Patent No. 6592877  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 355  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; City: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,967

; FILING DATE: 05-MAY-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Maki, David J.  
 ; REGISTRATION NUMBER: 31,392  
 ; REFERENCE/DOCKET NUMBER: 210121.411C9  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 107:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 391 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-09-072-967-107

Query Match 100.0%; Score 1949; DB 4; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-154;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60  
 DB 1 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60  
 QY 61 SSAGLWVAAAAPPYVAMNSVTAGQAEELTAAQVRVAAAAAYETAYGLTPPPVIAENRAELMI 120  
 DB 61 SSAGLWVAAAAPPYVAMNSVTAGQAEELTAAQVRVAAAAAYETAYGLTPPPVIAENRAELMI 120  
 QY 121 LIATNLLGONTPTAIVNEAEGYEMWQDAAMFGYAAATATATATLPPFEAPMTSAGG 180  
 DB 121 LIATNLLGONTPTAIVNEAEGYEMWQDAAMFGYAAATATATATLPPFEAPMTSAGG 180  
 QY 181 LLEQAAVEERASDTAAANQLMNNVPQALQAOPTQGTTPSSKLGGLWKTVPSPHSPISN 240  
 DB 181 LLEQAAVEERASDTAAANQLMNNVPQALQAOPTQGTTPSSKLGGLWKTVPSPHSPISN 240  
 QY 241 MVSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 300  
 DB 241 MVSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 300  
 QY 301 LGGGVAANLGRAASVGSLSVQPAWAAANQAVTPAARALPLTSLTSAABERGQMLGGLPV 360  
 DB 301 LGGGVAANLGRAASVGSLSVQPAWAAANQAVTPAARALPLTSLTSAABERGQMLGGLPV 360  
 QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391  
 DB 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391

## RESULT 6

; US-09-287-849-26  
 ; Sequence 26, Application US/09287849  
 ; Patent No. 6627198  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Dillon, David C.  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Campos-Neto, Antonio  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; TITLE OF INVENTION: and Their Uses  
 ; FILE REFERENCE: 014058-009020US  
 ; CURRENT APPLICATION NUMBER: US/09/287,849  
 ; CURRENT FILING DATE: 1999-04-07  
 ; PRIOR APPLICATION NUMBER: US 08/818,112  
 ; PRIOR FILING DATE: 1997-03-13  
 ; PRIOR APPLICATION NUMBER: US 08/942,578  
 ; PRIOR FILING DATE: 1997-10-01  
 ; PRIOR APPLICATION NUMBER: US 09/025,197  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 09/056,556

; PRIOR FILING DATE: 1998-04-07  
 ; PRIOR APPLICATION NUMBER: US 09/223,040  
 ; PRIOR FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 26  
 ; LENGTH: 596  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion  
 ; US-09-287-849-26

Query Match 100.0%; Score 1949; DB 4; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-154;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60  
 DB 9 MVDGALPPEINSARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 68  
 QY 61 SSAGLWVAAAAPPYVAMNSVTAGQAEELTAAQVRVAAAAAYETAYGLTPPPVIAENRAELMI 120  
 DB 69 SSAGLWVAAAAPPYVAMNSVTAGQAEELTAAQVRVAAAAAYETAYGLTPPPVIAENRAELMI 128  
 QY 121 LIATNLLGONTPTAIVNEAEGYEMWQDAAMFGYAAATATATATLPPFEAPMTSAGG 180  
 DB 129 LIATNLLGONTPTAIVNEAEGYEMWQDAAMFGYAAATATATATLPPFEAPMTSAGG 188  
 QY 181 LLEQAAVEERASDTAAANQLMNNVPQALQAOPTQGTTPSSKLGGLWKTVPSPHSPISN 240  
 DB 189 LLEQAAVEERASDTAAANQLMNNVPQALQAOPTQGTTPSSKLGGLWKTVPSPHSPISN 248  
 QY 241 MVSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 300  
 DB 249 MVSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 308  
 QY 301 LGGGVAANLGRAASVGSLSVQPAWAAANQAVTPAARALPLTSLTSAABERGQMLGGLPV 360  
 DB 309 LGGGVAANLGRAASVGSLSVQPAWAAANQAVTPAARALPLTSLTSAABERGQMLGGLPV 368  
 QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391  
 DB 369 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 399

## RESULT 7

; US-09-287-849-22  
 ; Sequence 22, Application US/09287849  
 ; Patent No. 6627198  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Dillon, David C.  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Campos-Neto, Antonio  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; TITLE OF INVENTION: and Their Uses  
 ; FILE REFERENCE: 014058-009020US  
 ; CURRENT APPLICATION NUMBER: US/09/287,849  
 ; CURRENT FILING DATE: 1999-04-07  
 ; PRIOR APPLICATION NUMBER: US 08/818,112  
 ; PRIOR FILING DATE: 1997-03-13  
 ; PRIOR APPLICATION NUMBER: US 08/942,578  
 ; PRIOR FILING DATE: 1997-10-01  
 ; PRIOR APPLICATION NUMBER: US 09/025,197  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 09/056,556  
 ; PRIOR FILING DATE: 1998-04-07  
 ; PRIOR APPLICATION NUMBER: US 09/223,040  
 ; PRIOR FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 46





Query Match 84.8%; Score 1652.5; DB 4; Length 396;  
Best Local Similarity 84.9%; Pred. No. 1.4e-129;  
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
DB 1 VUDFGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60

QY 61 SSAGLVAAASPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLVAAASPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLGQNTPAIAVNEAEYGENWADAAAFGYAAATATATATLLPPEEAPMTSAGG 180  
DB 121 LIATNLGQNTPAIAVNEAEYGENWADAAAFGYAAATATATATLLPPEEAPMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQOQAOTGTPSSKLGGLWKTVPSPHRSPLSN 240  
DB 181 LLEQAAVEEADTAAANQLMNNVPQALQOQAOTGTPSSKLGGLWKTVPSPHRSPLSN 240

QY 241 MVSMAHNSMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSS---LGSSL 296  
DB 241 IVSMLNHNVSMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSS---LGSSL 299

QY 297 GSSGLGGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERPGQMLG 356  
DB 300 GSSGLGAGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERPGHMLG 359

QY 357 GLPVGMGARAG--GGLSGVLRVPPRYVMPHSPAG 391  
DB 360 GLPLGQLTNSGGGFGVSNALRMPPRAYVMPRVPAAG 396

RESULT 12  
US-09-556-556-111  
; Sequence 111, Application US/09056556  
; Patent No. 6350456  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056,556  
; FILING DATE: 07-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.457  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 111:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 396 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-09-556-556-111

Query Match 84.8%; Score 1652.5; DB 4; Length 396;  
Best Local Similarity 84.9%; Pred. No. 1.4e-129;  
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60  
DB 1 VUDFGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60

QY 61 SSAGLVAAASPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120  
DB 61 SSAGLVAAASPYVAMSVTAGAELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLGQNTPAIAVNEAEYGENWADAAAFGYAAATATATATLLPPEEAPMTSAGG 180  
DB 121 LIATNLGQNTPAIAVNEAEYGENWADAAAFGYAAATATATATLLPPEEAPMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQOQAOTGTPSSKLGGLWKTVPSPHRSPLSN 240  
DB 181 LLEQAAVEEADTAAANQLMNNVPQALQOQAOTGTPSSKLGGLWKTVPSPHRSPLSN 240

QY 241 MVSMAHNSMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSS---LGSSL 296  
DB 241 IVSMLNHNVSMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQNGVRAMSS---LGSSL 299

QY 297 GSSGLGGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERPGQMLG 356  
DB 300 GSSGLGAGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERPGHMLG 359

QY 357 GLPVGMGARAG--GGLSGVLRVPPRYVMPHSPAG 391  
DB 360 GLPLGQLTNSGGGFGVSNALRMPPRAYVMPRVPAAG 396

RESULT 13  
US-09-072-596-106  
; Sequence 106, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel J.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,596  
; FILING DATE: 05-MAY-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031



Search completed: August 25, 2004, 03:14:26  
Job time : 36 secs

*This Page Blank (uspto)*